

DR WPI; 2000-23786

for m. g. m. s. i.


```
XX Page 1297; Disclosure; 1912pp; English.
PS
XX
CC Y34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 190 AA;

Query Match 38.4%; Score 363.5; DB 20; Length 190;
Best Local Similarity 39.7%; Pred. No. 6.2e-29;
Matches 73; Conservative 41; Mismatches 69; Indels 1; Gaps 1;

QY 2 ISVNDFKTGLTISVDNAIWKVIDFOHVKPGKGSFAVRSKLNRLRTGAIOEKTFRAGEKV 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 vstsefrvgrlieogqpyllqndfvpkgqgafnrkvnkftgrvrietyksgevs 63

QY 62 PAMIEENRMQYLYADGDNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIOTYEGETI 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64 tadiversmrillytdeqatfmddeffevevfweklenirqwlledtiytlvlyngdv 123

QY 122 GVELPKTVELVETEPGKIGDGTATG-ATKSATVETGYTLNVPLFVNEGDVLIINTGDGS 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
124 aveppifmelsaetapvrgdtasgrvlpavintgaklmvpifidegeivkvdttrgs 183

QY 181 YISR 184
Db : : : : :
184 yesr 187
```

```
XX Disclosure; Page 891; 1755pp; English.
PS
XX
CC Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome of
CC Chlamydia trachomatis (see Z01425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis, and
CC perihepatitis, Bartholin's disease, and lymphogranuloma venereum.
CC and venereal lymphogranulomatosis. The polypeptides of the invention
XX may be of use in treating these diseases.
XX
SQ Sequence 185 AA;

Query Match 29.6%; Score 280; DB 20; Length 185;
Best Local Similarity 31.7%; Pred. No. 1.4e-20;
Matches 59; Conservative 44; Mismatches 79; Indels 4; Gaps 2;

QY 1 MISVNDFKTGLTISVDNAIWKVIDFOHVKPGKGSFAVRSKLNRLRTGAIOEKTFRAGEKV 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 mvlssqslsvgnfistkdglykvsvskvsgnkgdtfkvsiqaagsdvivernfkaggev 60

QY 61 EPAMIEENRMQYLYADGDNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIOTYEGET 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 keafepnrileylyleedkylfldlgnvdkyikpkemknamfikaqvtvfvahvhegtv 120

QY 121 IGVELPKTVELVETEPGKIGD--TATGATKSATVETGYTLNVPLFVNEGDVLIINTGD 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 fsmelphfilelmvakt--fpgdsislsaggakkallegtvevlyppfveigdvikvdttr 178

QY 179 GSYISR 184
Db : : : : :
179 ceyigr 184

RESULT 5
Y34781
ID Y34781 standard; Protein; 185 AA.
XX
AC Y34781;
XX
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae transmembrane protein sequence.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST ) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Page 759-760; Disclosure; 1912pp; English.
```

CC *esp. E.coli* and *Pseudomonas fluorescens* for insecticidal applications.
 CC The toxins may be formulated with a phagostimulant or attractant for
 CC beetles.
 XX
 SQ Sequence 1174 AA;
 Query Match 9.4%; Score 88.5; DB 11; Length 1174;
 Best Local Similarity 30.9%; Pred. No. 2.9;
 Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps
 QY 42 RNLRTGAIQEKTFRAGEKVEPAMENRRNQYLYADGDNRHVMFNDSFEQTELSSDYLKEE 101
 11:1 11:1
 Db 1085 rnlrivedayemnttasvnykptveeer---vtdvqadhdhceydrvrvnrvpvaqvctke 1141

QY 102 LNI LR EGM EV Q I Q I E G E T I G V E L P R I V E L L I M E 1135
| | | : | | : | | | : | | : |
Db 1142 leyf pet d k v w i e i -- get e g k f i v d n v e l l i m e 1173

RESULT	7
P60243	
ID	P60243 standard; Protein; 2179 AA.
XX	
AC	P60243;
XX	
DT	07-AUG-1991 (first entry)

XX DE Sequence encoding the entire genomic RNA of human rhinovirus.
XX
KW Monoclonal antibody; MAB; HRV; vaccine; ss.

OS	Human rhinovirus.
XX	
FH	Location/Qualifiers

FT	Region	/label= VP4 structural protein 70..331	/label= VP2 structural protein 332..567
FT	Region		
FT	Region		

FT	Region	/label= VF	568..856
FT	Region	/label= VF	857..1002

FT	Region	/label= 5B protein
FT	Region	1003..1099
FT	FT	/label= 5B protein
FT	Region	1100..1429

FT	Region	/label= protein 1B
FT	Region	1430..1514
FT	Region	1515..1537

FT	Region	1538..1719	/label= Protease
FT	Region	1720..2179	

EP	
XX	XX
PN	EP169146-A.
XX	XX

22 JAN 1985.
17-JUL-1985; 85EP-0401465.
XX

PA (MERI) MERCK & CO INC.
XX
PR 20-JUL-1984; 84US-0632785.
PR 10 APR 1985; 8503 0721733

PI Colonne RJ, Mitzutani
XX
DR WPI; 1986-022809/04.

DR N-PSDB; N60194.

XX New DNA encoding the entire genomic RNA of human rhinovirus 14 -
PT and monoclonal antibodies which block attachment or neutralise
PT infectivity of rhinovirus.
XX Example 11; Page 22-39; 80pp; English.
XX Sequence may be used for the manufacture of hybridoma cells
CC expressing the HRV or fragments thereof. The fusion products may be
CC used in immunisation, or to raise MAb's for passive treatment of HRV
CC infection.
XX
SQ Sequence 2179 AA;

Query Match 9.2%; Score 87.5; DB 7; Length 2179;
Best Local Similarity 24.5%; Pred. No. 8.5;
Matches 39; Conservative 25; Mismatches 74; Indels 21; Gaps 7;
QY 26 QHVKPKGSFVRSKLNLRGTGAIQEKTFRAGEKVEPAMIEENRRMOYLYAGDGNH--VFM 83
Db 1713 qyfevkgqgvlarhkvrefninpvtat---ksklhpsvf-----ydvfpdgkepavls 1763
QY 84 DNSEPEQTELSGDLKEELNVLK----EGMEVQIQTYEGETIGVELPKTVELTVTETEPG 139
Db 1764 dndprlevkltleslskykgvntteptenmlvavdhyggllsidip-tseiltkealyg 1822
QY 140 IKGDTATGATKSAIVETGYTLNPLFVNEGDLVIINTGD 178
Db 1823 vdglepiddittsa----gfpv-vsigikkrdlmketqd 1856

RESULT 8
R25825
ID R25825 standard; Protein; 1174 AA.
XX
AC R25825;
XX
DT 26-JAN-1993 (first entry)
XX
DE Novel toxin expressed by PS81A2.
XX
KW Bacillus thuringiensis; lepidoptera; insect pest; insecticide; ss.
XX
OS Bacillus thuringiensis.
XX
PN AU9056291-A.
XX
PD 18-JUN-1992.
XX
PF 05-JUN-1990; 90AU-0056291.
XX
PR 05-JUN-1990; 90AU-0056291.
XX
PA (MYCO) MYCOGEN CORP.
XX
XX WPI; 1992-259364/32.
DR N-PSDB; Q26928.
XX
XX Controlling insect pests of lepidoptera family - includes
PT contacting insects with suitable amt. of Bacillus thuringiensis
XX
XX Claim 15; Page 29; 49pp; English.
XX
XX The protein sequence shows a Bacillus thuringiensis endotoxin whose
CC DNA may be used in an expression system to transform a variety of
CC microbial hosts, e.g. Pseudomonas, Azobacter, Erwinia, Serratia,
CC Agrobacterium, Streptomyces species, etc. Expression of the toxic
CC gene results in the intracellular prodn. and maintenance of the
CC peptide toxin. With suitable hosts, the microbes can be applied to
CC the sites of lepidopteran insects, e.g. to the insects themselves,
CC to the rhizosphere, phylloplane or to a body of water, where they
CC will proliferate and be ingested by the insects. The result is the

CC control of unwanted insects. Alternatively, the microbe hosting the
CC toxin gene can be treated under conditions that prolong the
CC activity of the toxin produced in the cell. The treated cell can
CC be applied to the environment of target pests. The resulting prod.
CC retains the toxicity of the B.t. toxin. See also R25826.
XX
SQ Sequence 1174 AA;

Query Match 9.1%; Score 86.5; DB 13; Length 1174;
Best Local Similarity 30.9%; Pred. No. 4.6;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;
QY 42 RNLRTGAIQEKTFRAGEKVEPAMIEENRRMOYLYAGDGNHVFMDNESFQTELSDDLKEE 101
Db 1085 nrnyedayemnttasvnykptyeere---ytdvggdhnceydrgvnyrvpvgayvtke 1141
QY 102 LNYLKEGNEVOIQTYEGETIGVELPKTVELTVTE 135
Db 1142 leyfpetdkvwlel--getegkfivdnvlellme 1173

RESULT 9
R89494
ID R89494 standard; Protein; 1174 AA.
XX
AC R89494;
XX
DT 28-MAY-1996 (first entry)
XX
XX B.t. toxin 81A2.
XX
KW Toxin; delta-endotoxin; Bacillus thuringiensis; crystal protein;
KW biological control; insecticide; pesticide; lepidoptera; CryIF;
KW insect resistance; transgenic plant.
XX
OS Bacillus thuringiensis isolate PS81A2.
XX
PN W09605314-A2.
XX
PD 22-FEB-1996.
XX
PF 14-AUG-1995; 95WO-US10310.
XX
PR 15-AUG-1994; 94US-0291368.
XX
PA (MYCO) MYCOGEN CORP.
XX
PI Narva KE, Payne J, Schnepf HE, Schwab GE, Sick AJ;
XX
XX WPI; 1996-139708/14.
DR N-PSDB; T16558.
XX
XX Toxin encoded by Bacillus thuringiensis PS91C2 nucleotide sequence -
PT and isolates and delta endotoxins obtd. from it, used to control
PT lepidopteran pests.
XX
PS Disclosure; Page 28-31; 52pp; English.
XX
XX Toxin 81A2 (R89494) is active against lepidopteran (caterpillar)
CC pests. It is the product of a gene (T16558) isolated from Bacillus
CC thuringiensis isolate PS81A2. B.t. isolates expressing CryIF class
CC toxin genes can be used for the biological control of lepidopteran
CC pests. Expression of the toxin by transgenic plants improves
CC resistance to such insects.
XX
SQ Sequence 1174 AA;

Query Match 9.1%; Score 86.5; DB 17; Length 1174;
Best Local Similarity 30.9%; Pred. No. 4.6;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;

CC 2 previously identified ABC drug efflux pumps. The predicted
 CC topology includes 12 transmembrane-spanning alpha-helices and 2
 CC highly conserved nucleotide binding domains. Constitutive
 CC overexpression of the PDR12 gene in yeast cells confers increased
 CC resistance to weak acids such as sorbate, propionate and benzoate,
 CC used e.g. as preservatives in foods and beverages. The invention
 CC provides a transformed yeast strain that constitutively expresses
 CC a gene encoding a weak acid pump. A vector comprising the PDR12
 CC gene and a process which comprises transforming or transfecting
 CC a yeast cell with the vector are claimed. The yeast cells can be
 CC used for the production of a dough or bread product (claimed).
 CC They can also be used for the production of an alcoholic beverage
 CC (e.g. whisky, wine, or beer) or other fermented product (claimed).
 CC Use of PDR12 as a weak acid pump is also claimed.

xx
 SQ Sequence 1511 AA;

Query Match 8.5%; Score 80.5; DB 20; Length 1511;
 Best Local Similarity 24.9%; Pred. No. 26;
 Matches 53; Conservative 31; Mismatches 80; Indels 49; Gaps 12;

Qy 12 TISVDNATWKVID--FQHVKPKGSAFV-----RSKLRNLTGAIQEKTFR----AGEK 59
 Db 850 tipdgatrklisdvfgvypgkmtalmgesagkttlinvlaqrimgvitgdmlvnak 909

Qy 60 VEPAMIENRRMQYLYADGDNHV--FMDNESF-----EQTESSDYLKELNYLKE-----G 108
 Db 910 plpasf-nrscgyv-aqadnmaelsvreslrfaaelrqgssvpleekyeyvekiitilg 967

Qy 109 MEVQIQTYEGET-----IGVELPKTVELVTETEP--GIKGDATGATK--S 151
 Db 968 mgnyaalvgktgrglnveqrkkisigvelvakpsllifldeptsldsgsawsivqfmr 1027

Qy 152 ATVETGYTL-----NVPLFVNEGDVLIINTG 177
 Db 1028 aladsggsilctihqpsatifeqforlillkkq 1060

Search completed: January 12, 2001, 14:58:32
 Job time: 36 sec

QY 01 SAEGADVDMNLIYLYNDGEFWHFMNNETFEQLSADAKAIGDNAKWLLDQAE CIVILWNG 120

Db 61 SAEGADVDDMLTYLYNDGEFWHFMNNTEFEQLSADAKAIGDNKWLDDOAEICVTIWLNG 120
Qy 121 QPISVTPNPFVELEIVTDPLGKGTAGTGKGPATLSTGAVVKVPLFVQIGEVIKVDTRS 180
Db 121 QPISVTPNPFVELEIVTDPLGKGTAGTGKGPATLSTGAVVKVPLFVQIGEVIKVDTRS 180
Qy 181 GEYVSRVK 188
Db 181 GEYVSRVK 188
RESULT 2
F82047
translation elongation factor EF-P VC2660 [imported] - Vibrio cholerae (group O1 strain
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: F82047
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, B.
l, R.C.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: F82047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <HEI>
A:Cross-references: GB:AE004332; GB:AE003852; NID:g9657252; PIDN:AAF95801.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2660
A:Map position: 1
C:Superfamily: translation elongation factor EF-P
Query Match 76.9%; Score 750; DB 2: Length 188;
Best Local Similarity 73.9%; Pred. No. 3.4e-60;
Matches 139; Conservative 26; Mismatches 23; Indels 0; Gaps 0;
Qy 1 MATYNSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRRLTGTREKTFKSTD 60
Db 1 MATYSTNEFGGLKIMLDNEPCVILEYVYKPGKGQAFNRIRKRLTGKVKLETKFSGD 60
Qy 61 SAEGADVDDMLTYLYNDGEFWHFMNNTEFEQLSADAKAIGDNKWLDDOAEICVTIWLNG 120
Db 61 TAEVADVDDIDLTYLYNDGEFWHFMNNSTFEQLAADAKAVGENAKWLVENNTCMLTLNG 120
Qy 121 QPISVTPNPFVELEIVTDPLGKGTAGTGKGPATLSTGAVVKVPLFVQIGEVIKVDTRS 180
Db 121 NPIAVTPNPFVELEIVTDPGKGTAGTGKGPATLSTGAVVRVPLFVQIGEVIKVDTRS 180
Qy 181 GEYVSRVK 188
Db 181 AEYVGRVK 188
RESULT 3
I64061
translation elongation factor EF-P - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C:Accession: I64061
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: I64061
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-188 <TIGR>

A:Cross-references: GB:U02717; GB:L42023; NID:g1573283; PIDN:AAC21989.1; PID:g1573295
C:Superfamily: translation elongation factor EF-P
C:Keywords: protein biosynthesis
Query Match 76.8%; Score 749; DB 2: Length 188;
Best Local Similarity 75.0%; Pred. No. 4.2e-60;
Matches 141; Conservative 20; Mismatches 27; Indels 0; Gaps 0;
Qy 1 MATYNSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRRLTGTREKTFKSTD 60
Db 1 MATYTTSDFPKGLKFMODGEPVIVENEFVKPGKGQAFTRTRIRKLISGKVLVDNFKSGT 60
Qy 61 SAEGADVDDMLTYLYNDGEFWHFMNNTEFEQLSADAKAIGDNKWLDDOAEICVTIWLNG 120
Db 61 SVEAADVMDLNLISYKDDAFWFHMPETFEQYSADAKAVGDAEKWLLDQADCIIVTLNG 120
Qy 121 OPTSVTPNPFVELEIVTDPLGKGTAGTGKGPATLSTGAVVKVPLFVQIGEVIKVDTRS 180
Db 121 APITVTPNPFVELEIVTDPLGKGTAGTGKGPATLSTGAVVKVPLFVQIGEVIRVDTRS 180
Qy 181 GEYVSRVK 188
Db 181 GEYVSRVK 188
RESULT 4
B82554
translation elongation factor EF-P XF2473 [imported] - Xylella fastidiosa (strain 9a5
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82554
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <SIM>
A:Cross-references: GB:AE004055; GB:AE003849; NID:g9107661; PIDN:AAF85271.1; GSPDB:GN
A:Experimental source: strain 9a5
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Saw
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2473
C:Superfamily: translation elongation factor EF-P
Query Match 65.4%; Score 638; DB 2: Length 193;
Best Local Similarity 63.3%; Pred. No. 4e-50;
Matches 119; Conservative 26; Mismatches 43; Indels 0; Gaps 0;
Qy 1 MATYNSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRRLTGTREKTFKSTD 60
Db 6 MASYGNNDVKNKGKILVNAEPVITDTEYVKPGKGQAFTRVKYRLIKSGRVQVETKSTD 65
Qy 61 SAEGADVDDMLTYLYNDGEFWHFMNNTEFEQLSADAKAIGDNKWLDDOAEICVTIWLNG 120
Db 66 TLEAADVDDTDMOYLYSDGEYWHFMQOETFEQVQADKNGMGGAEKWLKGEOCVVTLNG 125

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: B72021
A:Molecule type: DNA
A:Residues: 1-190 <ARN>
A:Cross-references: GB:AE001363; NID:94377212; PIDN:AAAD19033.1; PID:94377212
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255
A:Accession: B81517
A:Molecule type: DNA
A:Residues: 1-190 <REA>
A:Cross-references: GB:AE002256; GB:AE002161; NID:97189893; PIDN:AAF38751.1; PID:97189893
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: efp_2; CP0971
C:Superfamily: translation elongation factor EF-P

Query Match 45.7%; Score 445.5; DB 2; Length 190;
Best Local Similarity 46.3%; Pred. No. 7.5e-33;
Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;
QY 1 MATYNSNDFRAGIKIMLDGEPYAVEASEFVKPGKGOAFARVRLRLTGTVEKTFKSTD 60
Db 1 MVRVSTSEFRVGRUIDGQYLILQNDVFKPGKGOAFNRKFNLTGRVRIETYSKGE 60
QY 61 SAEGADVVDMLTYLYNDGEFHFMMNETFEQLSADAKAIGDNKAKWLLDQAEICVTIWLNG 120
Db 61 SVETADIVERSMRLTYDQCAFMDDETPEQEVFVEKLENIRQWLLDITVLYNG 120
QY 121 QPISVTPNPFVEIVDTPGLKGDTA-GTGGKRPATLSTGAVKVPFLFVQIGEVKVDTR 179
Db 121 DVAAVEPIEFMELSIETAEPVGRGDTASGRVLPKPAVNTGAKIMVPIFIDEGELVKVDTR 180
QY 180 SGEVVSRSV 187
Db 181 TGSYSRV 188

RESULT 9
G81401
translation elongation factor EF-P Cj0551 [similarity] - Campylobacter jejuni (strain NC
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: G81401
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barr
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912
A:Accession: G81401
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:96967817; PIDN:CA875187.1; PID:9696801
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: efp; Cj0551
C:Superfamily: translation elongation factor EF-P

Query Match 44.9%; Score 438; DB 2; Length 189;
Best Local Similarity 45.7%; Pred. No. 3.5e-32;
Matches 85; Conservative 33; Mismatches 68; Indels 0; Gaps 0;
QY 1 MATYNSNDFRAGIKIMLDGEPYAVEASEFVKPGKGOAFARVRLRLTGTVEKTFKSTD 60
Db 1 MASYSMDLKKGLKUIDGIPFKIVEYQHVKPGKGPAPFRIKISFDIGKVLKTFHAGD 60

QY 61 SAEGADVVDMLTYLYNDGEFHFMMNETFEQLSADAKAIGDNKAKWLLDQAEICVTIWLNG 120
Db 61 KCEAPNLEDKTMOYLYDDGNCQFMDTQTYEQVAISDDDDGEAKKWLMDGMWVDVLFHNG 120
QY 121 QPISVTPNPFVEIVDTPGLKGDTA-GTGGKRPATLSTGAVKVPFLFVQIGEVKVDTR 180
Db 121 KAIGVEVPQVVELKIETAPNFKGTQGSNKKPATLETGAVVQVFFHVEGEVIRVDTR 180
QY 181 GEYYSR 186
Db 181 GEYIER 186
RESULT 10
T30278
translation elongation factor EF-P - Synecchococcus sp. (PCC 7942)
C:Species: Synecchococcus sp.
A:Variety: PCC 7942
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C:Accession: T30278
R:Phung, L.T.; Haselkorn, R.
submitted to the EMBL Data Library, May 1996
A:Description: Genes encoding biotin carboxyl carrier protein and elongation factor P
A:Reference number: Z20804
A:Accession: T30278
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-185 <PHU>
A:Cross-references: EMBL:U59235; NID:gl399825; PID:gl399829; PIDN:AAB82025.1
C:Genetics:
A:Note: efp
C:Superfamily: translation elongation factor EF-P

Query Match 43.7%; Score 426; DB 2; Length 185;
Best Local Similarity 48.6%; Pred. No. 4.1e-31;
Matches 88; Conservative 25; Mismatches 68; Indels 0; Gaps 0;
QY 6 SNDFRAGIKIMLDGEPYAVEASEFVKPGKGOAFARVRLRLTGTVEKTFKSTDSEGA 65
Db 4 SNDFRTGTTIIDGAVWRVVEFLHVKPGKSAFVRTKLKNAKTGNVVKETFRAGETVPOA 63
QY 66 DVVDMLTYLYNDGEFHFMMNETFEQLSADAKAIGDNKAKWLLDQAEICVTIWLNGQPIVS 125
Db 64 VLEKSTLQYTKDGDVFDVMDMETYESRLTAATIGRVKYLKRGMEANVTWNGQVIEV 123
QY 126 TPNFVLEIVDTPGLKGDTA-GTGGKRPATLSTGAVKVPFLFVQIGEVKVDTRSGEYS 185
Db 124 ELPNSVLEIVETDPGVKGTATGTGTPAKVETGAQVMVPLFISVGERIKIDTRNDSYLG 183
QY 186 R 186
Db 184 R 184

RESULT 11
S74333
translation elongation factor EF-P - Synecchocystis sp. (strain PCC 6803)
N:Alternate names: hypothetical protein slr0434
C:Species: Synecchocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S74333
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocys
s.
A:Reference number: S74322; MUID:97061201
A:Accession: S74333
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-187 <KAN>

```
Db      1 MAT--TADFNGLVLDGQLWITTEQHVKPGKGPAFVRTKLKNVLSCKVVDKTFNAGV 58

Qy     61 SAEAGDVDMNLTYLYNDGEFFHFMMNETFEQLSADAKAIGNAKWLDDQACIVTLNNG 120
```

Db 59 KVDATVDRDRTTLYIRDGSDVFWDSDQIEQHPLPENALVGDAARFLLEGMPVQVAFIG 118

Qy 121 QPISVTPNPFVEIVDTDPGLKGDTAGTGCKPATLSTGAAYKVPFVQIGEVIKVDTRS 180
I : I : IIII : I : IIII : I : IIII : I : IIII :

Db 119 VPLYIELPVTVLELVTHTEPCLOGDRSSAGTKPAT'LOTGAQINVPFLINTGDKLKVDSDR 178

Qy 181 GEXYSRV 187
I : I : II

Db 179 GSYLGRV 185

RESULT 15
A9620 translation elongation factor EF-P efp - Bacillus subtilis
C:Species: Bacillus subtilis
C:date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: A9620
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Berteaux-Rabreau, C.; Bron, S.; Broutelle, S.; Bruschii, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, E.E.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997.
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallenbach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, V.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seronakeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, K.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zamanote, E.; Yoshikawa, H.; Danchin, A.
A:title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MID:98044033
A:Accession: A9620
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-185 <UN>
A:CROSS-references: GB:Z99116; GB:AL009126; NID:g2634723; PIDN:CAB14376.1; PID:g2634879
A:Experimental source: strain 168
C:Genetics:
C:Superfamily: translation elongation factor EF-P

	Query Match	39.0%;	Score 380;	DB 2;	Length 185;
	Best Local Similarity	43.3%;	Pred. No. 5.5e-27;		
	Matches 78;	Conservative 30;	Mismatches 72;	Indels 0;	Gaps
QY	7	NDFRAGLIKMLDGEPEVASEPVPKGGQAFARVKLRLLTGTRVEKTFKSTDSAGAD	66		
		: : : : : : : : : :			
Db	5	NDFRTGLTIDVCGGIWRKVDQFQVHPKGGAAFPVRSKLRLRTGATQEKTRFAGEKVAKAQ	64		
		: : : : : : : : : :			
QY	67	VDMNLTLYNDGEFPHFMNNETFEQLSADAKAIGDNAKWLLIDQAEICIVTLNNGQPISTV	126		
		: : : : : : : : : : : : : : : : : : : :			
Db	65	IETKTWOYLXANGDQHVEMDTSSYEQLSLSATQIEELKYLNNMVHMVQDETLGIT	124		
		: : : : : : : : : : : : : : : : : : : :			
QY	127	PPNFVELEIVDTPDLKGDGTAGTGGGKPAATLSTGAVVVKVPLFVIOGIVIKVDFRGGYVSR	186		
		: : : : : : : : : :			
Db	125	LPTNVLKVETEPGFKGDTSAGTGPARTETGLVYVNPFPFVNEGDTLVNNTSDGSYVSR	184		
		: : : : : : : : : :			

Search completed: January 12, 2001, 14:59:24
Job time: 84 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:58:00 ; Search time 24.94 seconds
(without alignments)
503.674 Million cell updates/sec

Title: US-09-322-732-1
Perfect score: 946
Sequence: 1 MISVNDFKTGLTISVDNAIW.....VNEGDLVLIITGDSYISRG 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	702	74.2	185	2 A69620 translation elonga
2	541	57.2	185	2 T30278 translation elonga
3	535	56.6	187	2 S74333 translation elonga
4	480	50.7	187	2 B70658 probable efp prote
5	443	46.8	185	2 D75558 translation elonga
6	435	46.0	192	2 E70418 translation elonga
7	434	45.9	189	2 G81401 translation elonga
8	421	44.5	185	2 B72212 translation elonga
9	409	43.2	187	2 A64542 translation elonga
10	406	42.9	187	2 D71967 translation elonga
11	383.5	40.5	190	2 E71475 probable translati
12	379.5	40.1	190	2 B81738 translation elonga
13	372	39.3	193	2 B82554 translation elonga
14	371	39.2	188	2 F82047 translation elonga
15	363.5	38.4	190	2 B72021 translation elonga
16	363	38.4	187	2 G71312 probable translati
17	362	38.3	188	2 S34443 translation elonga
18	343.5	36.3	188	2 G82908 translation elonga
19	337	35.6	190	2 H64202 translation elonga
20	331.5	35.0	188	2 B71678 translation elonga
21	327	34.6	190	2 S73451 translation elonga
22	311	32.9	188	2 I64061 translation elonga
23	291	30.8	185	2 B81708 translation elonga
24	290.5	30.7	188	2 F83290 translation elonga
25	284	30.0	275	2 B64986 hypothetical 30.9
26	275	29.1	185	2 F71553 probable translati
27	273.5	28.9	199	2 D82229 translation elonga
28	270	28.5	185	2 D72110 translation elonga
29	269.5	28.5	192	2 F70126 translation elonga

ALIGNMENTS

RESULT 1
A69620
translation elongation factor EF-P efp - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: A69620
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
A.; Fritschy, J.N.; Garavito, R.M.; Gots, S.; Gots, J.P.; Gots, S.; Gots, S.; Gots,
A.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033
A:Accession: A69620
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-185 <KUN>
A:Cross-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14376.1; PID:g26348
A:Experimental source: strain 168
C:Genetics:
A:Gene: efp
C:Superfamily: translation elongation factor EF-P

Query Match	74.2%	Score 702;	DB 2;	Length 185;
Best Local Similarity	71.7%	Pred. No. 1.1e-50;		
Matches 132;	Conservative 18;	Mismatches 34;	Indels 0;	Gaps 0;
Qy	1	MISVNDFKTGLTISVDNAIWVDFQHVKPGKGSFVRSKLNLTGAIQKTFRAGEKV	60	
Db	1	MISVNDFKTGLTISVDNAIWVDFQHVKPGKGSFVRSKLNLTGAIQKTFRAGEKV	60	
Qy	61	EPAMTENVRMQVLYADGNHFMVMDNESFEQTELSDYKLEELNKEGVEVOIQYEGET	120	
Db	61	AKAQETKTMQVLYANGQHFVMDTSSYEQLSELSATQIEELKYLENWSVIMMYQDET	120	
Qy	121	IGVELPKTVELTVEPCIKIDTATGATKTSATVETGYTLNPLFVNECDVLIINTGDS	180	
Db	121	LGIELPNTVELKVVETEPGKIDTASGCTKPAKTGLVNVVFPFVNEGDTLVVNTSDGS	180	
Qy	181	YISR 184		
Db	181	VYSR 184		

R. White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.R.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: D75558

C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316
 A:Accession: B72212
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-185 <ARN>
 A:Cross-references: GB:AE001815; GB:AE000512; NID:g4982341; PIDN:AAD36827.1; PID:g498
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TMI763
 C:Superfamily: translation elongation factor EF-P

Query Match 44.5%; Score 421; DB 2; Length 185;
 Best Local Similarity 44.6%; Pred. No. 1.3e-27;
 Matches 82; Conservative 35; Mismatches 67; Indels 0; Caps 0;
 Qy 1 MISVNDPKTGLTISVDNAINWKVIDFQHWKPGKGSFVRSKLRNLRTGAIGKTFRAGEKV 60

1 MIEVGDLKGMFIYDGEIYRVLEASAKHFMRGSGLIIRTKLNKVKTFGVRVFNPPSGEKV 60
 61 EPAMIENRNRWOYLADGDNHVFMDNESFEQTESSDLYLKEELNLYLKEGMEVOIQTYGET 120
 62 QEAEISFRKQAYLRDGDGHYFYFLLDDYEQYALSEEIEGDAKYLYLVENMEVDLYFHFGETP 120
 121 IGVLPKTVELTVETBPGIKGDTATGATKATSVETCTYTLNVPLFVNEGDVLIINTGDGS 180
 122 IGIELPTVTVETVETSPFKGDPVSGGKPAVLETGLKITVPYFIEVGDKRIKVDTRTGE 180
 181 YISR 184
 182 YVGR 184
 RESULT 9
 A64542
 translation elongation factor EF-P - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
 C:Accession: A64542
 R:Romb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A.; Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467

A;Accession: A04342
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-187 <TOM>
A;Cross-references: GB:AE000538; GB:AE000511; NID:g2313263; PIDN:AAD07247.1; PID:g2313266
C;Superfamily: translation elongation factor EF-P

Query Match	43.2%	Score	409;	DB 2;	Length	187;
Best Local Similarity	40.4%;	Pred. NO.	1.3e-26;			
Matches	74; Conservative	45; Mismatches	64; Indels	0; Gaps	0;	

QY	2	ISVNDFTGTLTISVDNAIWKFIDFQHVHKPKGGSAFVRSKLRLNLTGTAIQIEKTFRAGEKVE	61
Db	3	IGMSELKAGLKIELGGVPYRIVEYQHVKPKGAFAFRAKIKSFLDGDGVIEKTFHAGDKCE	62
QY	62	PAMTENRRMOYLXADGNHVMFMNESFEQTELSDDYLKEELNYLKEGMEVOIOTYEGETI	121
Db	63	EPNLIVEXTMOYLVHDGTYOFMDIESYEQIALNDSQVGGEASKWMLDGMQOVQLIHNDKAI	122
QY	122	GVELPKPVELTVTETEPGIGKDGTATGATKSATVETGYTLNVPLFVNESGDVLINTDGSY	181
Db	123	SVDVPQVVALKIVETAENFNKGDTSSASKKPATLETGAVQVPPHVLGEETIIKVNTETEY	182
QY	182	ISR	184
Db	183	LEK	185

RESULT 10
D71967
translation elongation factor EF-P - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 17-Mar-2000
C:Accession: D71967
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557
A:Accession: D71967
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-187 <ARN>
A:Cross-references: GB:AE001454; GB:AE001439; NID:g4154666; PIDN:AAD05735.1; PID:g4154666;
A:Experimental source: Strain J99
C:Genetics:
A:Gene: efp
C:Superfamily: translation elongation factor EF-P

Query Match	42.9%;	Score 406;	DB 2;	Length 187;
Best Local Similarity	40.4%;	Pred. NO. 2.3e-26;		
Matches 74;	Conservative 44;	Mismatches 65;	Indels 0;	Gaps 0;

[illegible]

RESULT 11
E71475
probable translation elongation factor EF-P - Chlamydia trachomatis (serotype D, strain E71475)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: E71475
R:Stephens, R. S.; Kalman, S.; Lammel, C. J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; MUID:99000809
A:Accession: E71475
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <ARN>
A:Cross-references: GB:AE001347; GB:AE001273; NID:g3329210; PIDN:AA68347.1; PID:g332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: efp_2
C:Superfamily: translation elongation factor EF-P

Query Match	40.5%	Score 383.5;	DB 2;	Length 190;
Best Local Similarity	41.3%;	Pred. No. 1.6e-24;		
Matches 76; Conservative	39;	Mismatches 68;	Indels 1;	Gaps 1;

[illegible]

RESULT 12
B81738

translation elongation factor P TC0133 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: B81738
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255
A:Accession: B81738
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <TET>
A:Cross-references: GB:AE002280; GB:AE002160; NID:G7190162; PIDN:AAF39011.1; PID:G719016
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
C:Superfamily: translation elongation factor EF-P

Query Match 40.1%; Score 379.5; DB 2; Length 190;
Best Local Similarity 40.8%; Pred. No. 3.5e-24;
Matches 75; Conservative 40; Mismatches 68; Indels 1; Gaps 1;
Qy 2 ISVDFKTKLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLRLNRGTGAIQEKTFRAGEKVE 61
Db 4 VSTSEFRVGLRVEDIDGQPYVILONDFVKPGKQAFNRVRLIKVNFLLTGRVIEKFSGESIE 63
Qy 62 PAMENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETI 121
Db 64 TADVREQQMLLYTQEGATFMDDETFEQELIFWDLKLENIRQWLLDITVYLVRNGDVI 123
Qy 122 GVELPKTVELTETEPGKIGDTATG-ATKSATVETGYTLNPLFVNEGDVLIINTGDS 180
Db 124 SVEPIFIELSIAETAPGRVGTASGRVLKPAITTTGAKIMVPFIEEGEVKVDTRGS 183
Qy 181 YISR 184
Db 184 YESR 187

RESULT 13
B82554
translation elongation factor EF-P XF2473 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82554
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <SIM>
A:Cross-references: GB:AE004055; GB:AE003849; NID:G9107661; PIDN:AAF85271.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovsky-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:

Query Match 39.2%; Score 371; DB 2; Length 188;
Best Local Similarity 41.0%; Pred. No. 1.7e-23;
Matches 75; Conservative 35; Mismatches 73; Indels 0; Gaps 0;
Qy 2 ISVDFKTKLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLRLNRGTGAIQEKTFRAGEKVE 61
Db 4 VSTNEFKGLKIMLDNEPCVILENEVYKPGKQAFNRVRLIKVLTGKLVLETKFSGDTAE 63
Qy 62 PAMENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETI 121
Db 64 VADVVDIDLDLYNDGEEYHFMNNTSTFQLAADAKAVGENAKWLVENNTCMLTLWNGNPI 123
Qy 122 GVLPKTVELTETEPGKIGDTATGATVETGYTLNPLFVNEGDVLIINTGDSY 181
Db 124 AVTPNPFVELEVETEDPGVKGDGTQGTGKPKATLSLGAVVRVPLFVQIGEVIKVDTRSAY 183
Qy 182 ISR 184
Db 184 VGR 186

RESULT 15
B72021
translation elongation factor p CP0971 [imported] - Chlamydia pneumoniae (strains
A:Gene: XF2473
C:Superfamily: translation elongation factor EF-P
Query Match 39.3%; Score 372; DB 2; Length 193;
Best Local Similarity 40.9%; Pred. No. 1.5e-23;
Matches 74; Conservative 40; Mismatches 67; Indels 0; Gaps 0;
Qy 4 VNDFKTKLTISVDNAIKWKVIDFQHVKPGKSAFVRSKLRLNRGTGAIQEKTFRAGEKVEPA 63
Db 11 MNDVKNMGKILVNAEPAVITTEYVVKPGKQAFTRVKYRLIKSGRVQEVTKSTDTLEAA 70
Qy 64 MIENRRMOYLYADGDHNVFMNDSFEQTELSDDYLKEELNLYLKEGMEVQIQTEGETIGV 123
Db 71 DVVDTDMOYLYSDGEYWHFMQOETFEQVQADKNGMGAEKWLKGEQCVVTLWNGVPIGV 130
Qy 124 ELPKTVELTETEPGKIGDTATGATKATVETGYTLNPLFVNEGDVLIINTGDSYIS 183
Db 131 QPPNFEVLKITETDPLRGDTSGGGKPKATLETGAVVRVPLFVNODEVIKVDTRSGEYVS 190
Qy 184 R 184
Db 191 R 191

RESULT 14
F82047
translation elongation factor EF-P VC3660 [imported] - Vibrio cholerae (group O1 stra
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: F82047
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: F82047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <HEI>
A:Cross-references: GB:AE004332; GB:AE003852; NID:G9657252; PIDN:AAF95801.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Map position: 1
C:Superfamily: translation elongation factor EF-P

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:57:56 ; Search time 21.21 Seconds
(without alignments)
156.627 Million cell updates/sec

Title: US-09-322-732-1

Perfect score: 946

Sequence: 1 MISVNDFKTGLTISVDNAIW.....VNEGDLVIINTGDGVSIRG 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/1aa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.5	9.1	1174	1	US-08-040-751-3
2	86.5	9.1	1174	1	US-08-291-368-2
3	86.5	9.1	1174	2	US-08-962-190-2
4	86.5	9.1	1174	4	PCT-US95-10310-2
5	86.5	9.1	1174	5	5164180-4
6	84	8.9	2089	1	US-08-418-893D-23
7	84	8.9	2089	1	US-08-418-893D-24
8	83	8.8	294	1	US-08-137-175A-9
9	83	8.8	294	3	US-08-479-017-9
10	80	8.5	1430	3	US-09-008-172-2
11	78	8.2	257	1	US-07-781-355-2
12	78	8.2	2366	1	US-08-480-604A-10
13	78	8.2	2366	2	US-08-405-496A-10
14	77.5	8.2	464	2	US-09-021-323-3
15	76	8.0	663	3	US-08-776-265-5
16	76	8.0	770	1	US-08-525-654A-1
17	76	8.0	771	1	US-08-525-654A-3
18	75.5	8.0	1167	1	US-08-100-709-2
19	75.5	8.0	1167	1	US-08-176-865-2
20	75.5	8.0	1167	1	US-08-474-038-2
21	75.5	8.0	1167	2	US-08-779-046-2
22	75.5	8.0	1167	2	US-08-881-340-2
23	74.5	7.9	801	3	US-09-104-070-2
24	74.5	7.9	1176	1	US-08-434-823-2
25	74.5	7.9	1176	1	US-08-457-366-2
26	74.5	7.9	1535	3	US-08-755-587-185
27	74	7.8	693	1	US-08-553-279-2
28	73	7.7	526	2	US-08-853-659A-40

29 73 7.7 999 2 US-08-770-301A-1 Sequence 1, Appli
30 73 7.7 999 3 US-09-175-581-1 Sequence 1, Appli
31 72 7.6 644 1 US-08-487-890A-6 Sequence 6, Appli
32 72 7.6 644 2 US-08-478-435-6 Sequence 6, Appli
33 72 7.6 644 2 US-08-337-483-6 Sequence 6, Appli
34 72 7.6 644 2 US-08-478-373-6 Sequence 6, Appli
35 72 7.6 644 3 US-08-474-671-6 Sequence 6, Appli
36 72 7.6 644 3 US-08-483-577A-6 Sequence 6, Appli
37 72 7.6 664 3 US-08-669-408B-2 Sequence 2, Appli
38 72 7.6 1683 3 US-08-755-587-183 Sequence 183, App
39 71 7.5 277 1 US-08-400-413-1 Sequence 1, Appli
40 71 7.5 348 2 US-08-844-153-2 Sequence 2, Appli
41 71 7.5 419 3 US-09-155-200-4 Sequence 2, Appli
42 71 7.5 480 2 US-08-962-203-2 Sequence 2, Appli
43 71 7.5 480 3 US-09-282-125A-2 Sequence 2, Appli
44 71 7.5 500 1 US-08-260-582-77 Sequence 77, Appli
45 71 7.5 500 4 PCT-US95-05471-77 Sequence 77, Appli

ALIGNMENTS

RESULT 1
US-08-040-751-3
; Sequence 3, Application US/08040751
; Patent No. 5407825
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5407825el Bacillus thuringiensis isolates
; TITLE OF INVENTION: active against Lepidopteran Pests and Genes Encoding No. 54
; TITLE OF INVENTION: Lepidopteran-active Toxins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID R. SALIWANCHIK
; STREET: 2421 N.W. 41st STREET, SUITE A-1
; CITY: GAINESVILLE
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/040,751
; APPLICATION NUMBER: 19930329
; FILING DATE: 19930329
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SALIWANCHIK, DAVID R.
; REFERENCE/DOCKET NUMBER: MA39.C1.D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2

US-08-040-751-3

```

Query Match          9.1%; Score 86.5; DB 1; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;

QY  42  RNLGTAIQIKTFRAGEKVEPAMTENRRMQLYADGDNHVFMDNESFQETLSSDYLKEE 101
      |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1085  RNRIYEDAYENMTTASVNYKFTYEER---YTDVQGDHCEYDRGYVNRVPAGYVTK 1141
      |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY  102  LNYLKEGMEVOIQTYEGTIGVLPKTVLTVTE 135
      |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1142  LEYFETDKVWIEI--GETEGKFIVDNVELLME 1173
      |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-08-291-368-2
; Sequence 2, Application US/08291368
; Patent No. 5686069
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: No. 5686069el Bacillus thuringiensis Isolates
; TITLE OF INVENTION: Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/291,368
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2
; US-08-291-368-2

Query Match          9.1%; Score 86.5; DB 1; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;

```

```
Matches      29; Conservative   10; Mismatches    50; Indels       5; Gaps

Qy  42  KRLRTGAIOEKTFRAGEKVEPAMENRRMOYLYADGDHNVFMDNESFEOTELSSDYLKEE 101
     ||| | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db  1085 RNIRYEDAYEMNTTASVNYKPTVEER---YTDVGDNHCEYDRGYNVRPVPGYYTKE 114.1

Qy  102  LNYLKESGMEVOIQTYEGETIGVELPKTVTELTVTE 135
     | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db  1142 LEVPETDKWIEI--GETEGKFIVDNVELLME 1173

RESULT      3
US-08-962-190-2
; Sequence 2, Application US/08962190
; Patent No. 5985267
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: No. 5985267el Bacillus thuringiensis Isolates
; TITLE OF INVENTION: Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/962,190
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/291,368
; FILING DATE:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick.
; CLONE: 81A2
; US-08-962-190-2

Query Match          9.1%; Score 86.5; DB 2; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches      29; Conservative   10; Mismatches    50; Indels       5; Gaps

Qy  42  KRLRTGAIOEKTFRAGEKVEPAMENRRMOYLYADGDHNVFMDNESFEOTELSSDYLKEE 101
```



```

Db 1085 RNIRYEDAYEMNTTASVNYKPTYEEER---YTDVQGDHNHCEYDRGYVNYRVPVAGYVTKE 1141
Qy 102 LNYLKEGMEVOIQTYEGETIGVGLPKTVELTVTE 135
Db 1142 LEYFPETDKVWIEI--GETEGKFIVDNVELLME 1173

RESULT 4
PCT-US95-10310-2
; Sequence 2, Application PC/TUS9510310
; GENERAL INFORMATION:
; APPLICANT: MYCOGEN CORPORATION
; APPLICANT: STREET ADDRESS: 5501 Oberlin Drive
; APPLICANT: CITY: San Diego
; APPLICANT: STATE/PROVINCE: California
; APPLICANT: COUNTRY: US
; APPLICANT: POSTAL CODE/ZIP: 92121
; APPLICANT: PHONE NUMBER: (619) 453-8030
; APPLICANT: FAX NUMBER: (619) 453-6991
; TITLE OF INVENTION: Protein Toxins Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10310
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem - 11 (tm) Library of August Sick
; CLONE: 81A2
PCT-US95-10310-2
```

```

Query Match 9.1%; Score 86.5; DB 4; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;
Qy 42 RNLRTGATQKTFRAGEKVEPAMENRRMQYLYADGDNHVFMDNESFQTELSSDYLKEE 101
```

```

Db 1085 RNIRYEDAYEMNTTASVNYKPTYEEER---YTDVQGDHNHCEYDRGYVNYRVPVAGYVTKE 1141
Qy 102 LNYLKEGMEVOIQTYEGETIGVGLPKTVELTVTE 135
Db 1142 LEYFPETDKVWIEI--GETEGKFIVDNVELLME 1173

RESULT 5
5164180-4
; Patent No. 5164180
; APPLICANT: Payne, Jewel; Sick, August J.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS ISOLATES ACTIVE
; AGAINST LEPIDOPTERAN PESTS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/451,389
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 353,860
; FILING DATE: 18-MAY-1989
; SEQ ID NO:4:
; LENGTH: 1174
5164180-4

Query Match 9.1%; Score 86.5; DB 5; Length 1174;
Best Local Similarity 30.9%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 5; Gaps 2;
Qy 42 RNLRTGATQKTFRAGEKVEPAMENRRMQYLYADGDNHVFMDNESFQTELSSDYLKEE 101
Db 1085 RNIRYEDAYEMNTTASVNYKPTYEEER---YTDVQGDHNHCEYDRGYVNYRVPVAGYVTKE 1141
Qy 102 LNYLKEGMEVOIQTYEGETIGVGLPKTVELTVTE 135
Db 1142 LEYFPETDKVWIEI--GETEGKFIVDNVELLME 1173
```

```

RESULT 6
US-08-418-893D-23
; Sequence 23, Application US/08418893D
; Patent No. 5559220
; GENERAL INFORMATION:
; APPLICANT: ROESSLER, PAUL G
; APPLICANT: OHLROGGE, JOHN B
; TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
; TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTHELLA CRYPTICA
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
; STREET: 1617 Cole Blvd.
; CITY: Golden
; STATE: CO
; COUNTRY: USA
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,893D
; FILING DATE: April 7, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,938
; FILING DATE: September 14, 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: O'CONNOR, EDNA
; REGISTRATION NUMBER: 29,252
; REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON
```

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 303-231-1000
 TELEFAX: 303-231-1098
 TELEX:
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2089 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-418-893D-23

Query Match 8.98; Score 84; DB 1; Length 2089;
 Best Local Similarity 26.08; Pred. No. 4.5;
 Matches 53; Conservative 27; Mismatches 74; Indels 50; Gaps 13;
 QY 2 ISVNDFTGLTISVDNAI-----WK---VIDFQHVKPGKGSFV--RSKLRNLRGTGAI 49
 Db 261 IPMEIFNKLGLTSADEAVIVANKIGWENGIMIKASEGGGKIRFVDNEADLRNFAVQVS 320
 QY 50 QE-----KTRF-----AGEKVEPAMTEN-----RRMOYLYADGDNHVFMD 84
 Db 321 NEVIGSPIFLMQLCKNARHIEVQIVGDHGNVAVALNGRDCSTQRRFOKIFEEGPPSI-VP 379
 QY 85 NESFEQTELSDDYLKEELNYLKEG-MEVQIQTYEGETIGVEL-PK-TVELTVTETEPGAIK 141
 Db 380 KETHEMELAAORLTONTIGYOGACTVEYLYNAADNKKFFLELNPRLQVHPVTE---GI- 435
 QY 142 GDTATGATKATVETGYTLNPLF 165
 Db 436 ----TGANLPAT-QLQVAMGIPLF 454

RESULT 7
 US-08-418-893D-24
 Sequence 24, Application US/08418893D
 Patent No. 559220
 GENERAL INFORMATION:
 APPLICANT: ROESSLER, PAUL G
 APPLICANT: OHLSGEGE, JOHN B
 TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
 TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTOLLA CRYPTICA
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
 STREET: 1617 Cole Blvd.
 CITY: Golden
 STATE: CO
 COUNTRY: USA
 ZIP: 80401-3393
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/418,893D
 FILING DATE: April 7, 1995
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/104,938
 FILING DATE: September 14, 1993
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: O'CONNOR, EDNA
 REGISTRATION NUMBER: 29,252
 REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 303-231-1000

TELEFAX: 303-231-1098
 TELEX:
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2089 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 US-08-418-893D-24

Query Match 8.98; Score 84; DB 1; Length 2089;
 Best Local Similarity 26.08; Pred. No. 4.5;
 Matches 53; Conservative 27; Mismatches 74; Indels 50; Gaps 13;
 QY 2 ISVNDFTGLTISVDNAI-----WK---VIDFQHVKPGKGSFV--RSKLRNLRGTGAI 49
 Db 261 IPMEIFNKLGLTSADEAVIVANKIGWENGIMIKASEGGGKIRFVDNEADLRNFAVQVS 320
 QY 50 QE-----KTRF-----AGEKVEPAMTEN-----RRMOYLYADGDNHVFMD 84
 Db 321 NEVIGSPIFLMQLCKNARHIEVQIVGDHGNVAVALNGRDCSTQRRFOKIFEEGPPSI-VP 379
 QY 85 NESFEQTELSDDYLKEELNYLKEG-MEVQIQTYEGETIGVEL-PK-TVELTVTETEPGAIK 141
 Db 380 KETHEMELAAORLTONTIGYOGACTVEYLYNAADNKKFFLELNPRLQVHPVTE---GI- 435
 QY 142 GDTATGATKATVETGYTLNPLF 165
 Db 436 ----TGANLPAT-QLQVAMGIPLF 454

RESULT 8
 US-08-137-175A-9
 Sequence 9, Application US/08137175A
 Patent No. 5777095
 GENERAL INFORMATION:
 APPLICANT: BARBOUR, Alan G.
 APPLICANT: BERGSTROM, Sven
 APPLICANT: HANSSON, Lennart
 TITLE OF INVENTION: IMPROVEMENT IN BORRELIA BURGDORFERI AND
 TITLE OF INVENTION: PROPHYLAXIS
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROWDY AND NETMARK
 STREET: 419 Seventh Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/137,175A
 FILING DATE: 26-OCT-1993
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/08972
 FILING DATE: 22-OCT-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: COOPER, Iver P.
 REGISTRATION NUMBER: 28,005
 REFERENCE/DOCKET NUMBER: BARBOUR-1B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528

```
;
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-137-175A-9

Query Match      8.8%  Score 83; DB 1; Length 294;
Best Local Similarity 24.08; Pred. No. 0.32;
Matches 47; Conservative 30; Mismatches 61; Indels 58; Gaps 11;

Qy 1 MISVNDKTKTGLTISVDNAIKWVIDFQHVKPGKGSFVRSKLRNLTGAIQKTFRAGEKV 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 MLVSDLLNT-ITITYDPSNKKISSQVAK-----KQGSLTEETTKT-SKL 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 61 EPAMIEENRMQVLYADGNHVFMDNESFEQTELS-SDYLKEELNVLKRGMEVQIQTYEGE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 SAKKITR-----SNTTIEYTEMTDADNASKAVETLNGI-----TLEGS 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 120 TIGVELPKTV-ELVT-----ETEPGK---GDTATGATKSATV--ETGYTLNVP----- 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 LVGGKTTLTIKEGVTTLKKEIEKAGTVKFLDLDTTASSATKKTAVWNDTSSLTLSAEGKK 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 164 ----LFVNEGDVLIIN 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 TKDFVFLTDGTTIVQN 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-08-479-017-9
; Sequence 9, Application US/08479017
; Patent No. 6143872
; GENERAL INFORMATION:
; APPLICANT: BARBOUR, Alan G.
; APPLICANT: BERGSTROM, Sven
; APPLICANT: HANSSON, Leunart
; TITLE OF INVENTION: IMPROVEMENT IN BORRELIA BURGDORFERI AND
; TITLE OF INVENTION: PROPHYLAXIS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,017
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,175
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: PCT/US92/08972
; FILING DATE: 22-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: BARBOUR-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-479-017-9

Query Match      8.8%  Score 83; DB 3; Length 294;
Best Local Similarity 24.08; Pred. No. 0.32;
Matches 47; Conservative 30; Mismatches 61; Indels 58; Gaps 11;

Qy 1 MISVNDKTKTGLTISVDNAIKWVIDFQHVKPGKGSFVRSKLRNLTGAIQKTFRAGEKV 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 MLVSDLLNT-ITITYDPSNKKISSQVAK-----KQGSLTEETTKT-SKL 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 61 EPAMIEENRMQVLYADGNHVFMDNESFEQTELS-SDYLKEELNVLKRGMEVQIQTYEGE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 SAKKITR-----SNTTIEYTEMTDADNASKAVETLNGI-----TLEGS 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 120 TIGVELPKTV-ELVT-----ETEPGK---GDTATGATKSATV--ETGYTLNVP----- 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 LVGGKTTLTIKEGVTTLKKEIEKAGTVKFLDLDTTASSATKKTAVWNDTSSLTLSAEGKK 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 164 ----LFVNEGDVLIIN 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 TKDFVFLTDGTTIVQN 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nicholls, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-008-172-2

Query Match      8.5%  Score 80; DB 3; Length 1430;
Best Local Similarity 22.28; Pred. No. 7;
Matches 40; Conservative 31; Mismatches 87; Indels 22; Gaps 6;

Qy 15 VDNAIKWVIDFQHVKPGKGSFVRSKLRNLTGAIQKTFRAGEK-----VEPAMI 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 587 VQTVIAKIIKAQINPKTDGLFTLDELK--QAFKIYNEDMRQAKKYYTQSNIPATAYALML 644
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 66 ENR----RMQY--LYADGNHVFMDNESFEQTELSDDLKEELNVLKRGMEVQIQTYEGE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 645 SNKDSITRLYYGDMYSDGQYMATKSPYYDAID---TLLKARIKYAAGQDMKITYVEGD 701
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 120 TIGVELPKTVELTETEPGKIGDTATG--ATKSATVETGYTLNPLFVNEGDVLIINTG 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 702 KSHMDWDYTGVLTSVRYGTGANEATDQGEATKTQGMAVITSNPNPSLKLQNDKVINWNG 761
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-07-781-355-2
; Sequence 2, Application US/07781355
; Patent No. 5246844
; GENERAL INFORMATION:
; APPLICANT: No. 5246844ris, Steven J.
```

APPLICANT: Barbour, Alan G.
TITLE OF INVENTION: VIRULENCE ASSOCIATED PROTEINS IN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/781,355
FILING DATE: 19911022
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: USH:162
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-781-355-2

Query Match 8.2%; Score 78; DB 1; Length 257;
Best Local Similarity 24.2%; Pred. No. 0.92;
Matches 39; Conservative 24; Mismatches 64; Indels 34; Gaps 7;

Qy 23 IDFOHVKPGKSAFVRSKLRNLTGAIOEKTFRAGEKVEPAMENRRMOYLYADGDNHVF 82
Db 124 IDLEKIKNSDKAIFASML-----AKEAYSLTKAAEONMOKLY----- 161
Qy 83 MDNESFEQTESSDYL-KEELNYLKEGMEVQIQ-TYEGETIGVELPKTVLTETEPGI 140
Db 162 KEQOKISESESDYSDSAEIKQAKEAVEIAKATVEAKDKLID----VENTVKTLDKI 217
Qy 141 KGDRATGATKATVETGTYTLNPLFVNGDVL-----IINT 176
Db 218 KTET-TNNTKLADIKAAELVLQIAKNAKEIVQEVVALLNT 257

RESULT 12
US-08-480-604A-10
Sequence 10, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2366 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-10

Query Match 8.2%; Score 78; DB 1; Length 2366;
Best Local Similarity 21.5%; Pred. No. 24;
Matches 50; Conservative 39; Mismatches 74; Indels 70; Gaps 13;

Qy 1 MISVNDPK-----TGLTISVDNATWKVI---DFQHVKPGKGSFVRSK--- 40
Db 736 IVSANQVEVRINSEGRRELLDHSGEWINKESIIKDISKEYISFNPKENKITVKSNNLP 795
Qy 41 -----LRNLRGTGA-----IOEKT-----FRAGEKVEPAMENR---RMQYLYADGDNHV 81
Db 796 ELSTLLQEIENNNSSDIELEEKVMLTECEINVISNIDTQIVERRIEEAKNLTSDSINYI 855
Qy 82 FMDNESFEQTESSDYL-----KEELNYLKEGMEVQIQ-TYEGETIGVE 124
Db 856 ---KDEPKLIESIDALCDLKQONELEDHSFISFEDISETDEGSIFRINKEGESIFVE 912
Qy 125 LPKTV-----ELTVTEPGIKG---DPATG-ATKSATVETGY---TLNPLFV 166
Db 913 TEKTIFSEYANHTEIISKIKGTIFDTVNGKLVKKNLDTTHEVNTLNAAFFI 965

RESULT 13
US-08-405-496A-10
Sequence 10, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,496A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-405-496A-10

Query Match 8.2%; Score 78; DB 2; Length 2366;
Best Local Similarity 21.5%; Pred. No. 24;
Matches 50; Conservative 39; Mismatches 74; Indels 70; Gaps 13;

Qy 1 MISVNDKF-----TGLTISVDNAIWKVI---DFQHKVPGKGSFVRSK--- 40
Db 736 IVSAQYEVRIINSEGRRELLDHSGEWINKESIINKDISKEYISFPNKENKITVKSNNLP 795
Qy 41 -----LRLNRGA-----IOEKT-----FRAGEKVEPAMIEKN--RMOYLYADGDNHV 81
Db 796 ELSTLLQIRNNSSDIELEKVMLTECEINVISNIDTQVEERIEAKNLTSDSINIY 855
Qy 82 FDNESFEQTELSDDL-----KEELNLYKGEVQ-IQTYEGTIGVE 124
Db 856 ---KDEFKLIESIDALCDLQKQNELEDSHFISFEDSETDEGFSIRFINKETGESIFVE 912
Qy 125 LPKTV-----ELTVIETEGIGK---DTATG-ATKSATVETGY---TLNVPFLV 166
Db 913 TEKTIFSEYANHITEISKIRGTFIDTVNGKLVKVNLDTTHEVNTLNAAFFI 965

RESULT 14
US-09-021-323-3
; Sequence 3, Application US/09021323
; Patent No. 5929033
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry

; TITLE OF INVENTION: EXTRACELLULAR MUCOUS MATRIX
; TITLE OF INVENTION: GLYCOPROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,323
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0477 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 294502
; US-09-021-323-3

Query Match 8.2%; Score 77.5; DB 2; Length 464;
Best Local Similarity 22.3%; Pred. No. 2.5;
Matches 35; Conservative 22; Mismatches 45; Indels 55; Gaps 6;

Qy 9 TGLTISVDNAIWKVIDFOHKVPGKGSFVRSKLRLNRGTGAIOEKTFRAGEKVEPAMIE-- 66
Db 23 TGLAGDKHCVCVELLPDSSFPK-----RVGALEDETIRLSNRVEDEMOKLE 70
Qy 67 -----NRRMOYLYADGDNHVFMNDSFEQTELSDDLKEELNLYK---- 106
Db 71 EQDITLDITYSEKIINLTRLRVEYLEK-----LHPESL--VEISFEVLKREIRELEMYIS 121
Qy 107 -----EGMEVQIQTYEGTIGVLPKTVLTVTETE 137
Db 122 AMRVKPNNGSVQVETLYNEV-----KNMSKTVGQLE 152

RESULT 15
US-08-776-265-5
; Sequence 5, Application US/08776265
; Patent No. 6001631
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

```

; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/776,265
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 663 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-776-265-5

```

```

Query Match      8.0%; Score 76; DB 3; Length 663;
Best Local Similarity 26.2%; Pred. No. 6.1;
Matches 33; Conservative 22; Mismatches 29; Indels 42; Gaps 8;

Qy  8 KTG--LTISVDNAIKV---IDFQHVKPG-KGSAPVRSKLRNLRGTGAIQKTFRAGEKVE 61
Db 163 KTGKTVTFKPDITFKASTSFNFDVLSERLQESAF---LKNLK---ITLNDLRSGK-- 214

Qy  62 PAMIEENRMQVLYADGNHVPMDNESFEQTELSDDLKEELNLYLKEGMEV--QIQTYEGE 119
Db 215 -----RQEHYHBEG-----IKFVSYVNEGKVLHLDVATFSGE 248

Qy 120 TIGVEL 125
Db 249 ANGIEV 254

```

Search completed: January 12, 2001, 14:58:57
Job time: 61 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:58:40 ; Search time 16.19 Seconds
(without alignments)
369.018 Million cell updates/sec

Title: US-09-322-732-1

Perfect score: 946

Sequence: 1 MISVNDFKTGLTISVDNAIW.....VNEGDLVLIINTGDSYISRG 185

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	702	74.2	185	1 EFP_BACSU	P49778 bacillus su
2	541	57.2	185	1 EFP_SYNP7	Q54760 synechococ
3	535	56.6	187	1 EFP_SYNP3	Q55119 synechocyst
4	505	53.4	185	1 EFP_ANASP	Q44247 anabaena sp
5	480	50.7	187	1 EFP_MYCTU	P95019 mycobacteri
6	460	48.6	187	1 EFP_BRELA	Q45288 brevibacter
7	435	46.0	192	1 EFP_AQUAE	Q67376 aquifex aeo
8	421	44.5	185	1 EFP_THEMEA	Q9x284 thermotoga
9	409	43.2	187	1 EFP_HELPY	P36004 helicobacte
10	407	43.0	185	1 EFP_BACFR	P70889 bacteroides
11	406	42.9	187	1 EFP_HELPJ	Q92mq5 helicobacte
12	383.5	40.5	190	1 EFP2_CHLTR	Q84757 chlamydia t
13	363.5	38.4	190	1 EFP2_CHLPN	Q9x711 chlamydia p
14	363	38.4	187	1 EFP_TREPA	O83537 treponema p
15	362	38.3	187	1 EFP_ECOLI	P33398 escherichia
16	337	35.6	190	1 EFP_MYCGE	P47272 mycoplasma
17	331.5	35.0	188	1 EFP_RICPR	Q9zdt7 rickettsia
18	327	34.6	190	1 EFP_MYCPN	P75085 mycoplasma
19	311	32.9	187	1 EFP_HAEIN	P43771 haemophilus
20	284	30.0	275	1 YEIP_ECOLI	Q3028 escherichia
21	280.5	29.7	187	1 EFP_BUCMP	O51834 buchnera ap
22	275	29.1	185	1 EFP1_CHLTPN	Q44124 chlamydia t
23	270	28.5	185	1 EFP1_CHLTPN	Q9z900 chlamydia p
24	269.5	28.5	192	1 EFP_BORBU	O51232 borrelia bu
25	114.5	12.1	135	1 IF5A_SULAC	P28461 sulfolobus
26	104	11.0	132	1 IF5A_METJA	Q38625 methanococc
27	95.5	10.1	138	1 IF5A_PYRHO	O50089 pyrococcus
28	95	10.0	136	1 IF5A_PYRAE	P56635 pyrobaculum
29	94.5	10.0	386	1 DEGS_BACBR	P54663 bacillus br
30	91	9.6	280	1 DAPF_ARCFU	O29511 archaeoglob
31	91	9.6	5255	1 BKCA_BACLI	O68006 b bacitraci
32	90.5	9.6	203	1 CTC_BACSU	P14194 bacillus su
33	88	9.3	675	1 CIBA_PABPP	P57091 paenibacill

34 87 9.2 1951 1 CIN3_RAT P08104 rattus norv
35 86.5 9.1 130 1 IF5A_METH O26955 methanobact
36 86.5 9.1 639 1 GYRB_HALSQ P21558 haloferax s
37 86.5 9.1 1174 1 C1EB_BACTA Q03745 bacillus th
38 84.5 8.9 268 1 ENGA_SALTY Q9xc18 salmonella
39 83.5 8.8 364 1 YEJB_ECOLI P33914 escherichia
40 83.5 8.8 604 1 SP20_YEAST P50875 saccharomyc
41 83.5 8.8 2179 1 POLG_HRV14 P03303 human rhino
42 83 8.8 567 1 VE06_VARV P23819 variola vir
43 82 8.7 352 1 SISI_YEAST P25294 saccharomyc
44 82 8.7 567 1 VE06_VACCC P21047 vaccinia vi
45 82 8.7 567 1 VE06_VACCV P21607 vaccinia vi

ALIGNMENTS

RESULT 1
EFP_BACSU
ID EFP_BACSU STANDARD; PRT; 185 AA.
AC P49778;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,
RA Sato T., Takeuchi M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 139-185 FROM N.A.
RC STRAIN=168 / JH642;
RA Guerout-Fleury A.M., Gonzy-Treboul G., Stragier P.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN PEPTIDE-BOND SYNTHESIS. STIMULATE EFFICIENT
TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONING INDIRECTLY BY ALTERING
THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
(BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D84432; BAA12558.1; -;
CC EMBL; U35252; AAA76718.1; -;
CC EMBL; Z99116; CAB14376.1; -;
CC SUBTILIST; BG11460; EFP.
CC INTERPRO; IPR001059; -;
CC PFAM; PF01132; EFP; 1.
CC PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20454 MW; 60C3B9BF90D2DAB8 CRC64;

Query Match 74.2%; Score 702; DB 1; Length 185;
Best Local Similarity 71.7%; Pred. No. 1.2e-48;
Matches 132; Conservative 18; Mismatches 34; Indels 0; Gaps 0;


```

RESULT 4
EFP_ANASP
ID EFP_ANASP STANDARD; PRT; 185 AA.
AC Q44247;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
RN [1]
RX MEDLINE=93352435; PubMed=8102363;
RA Gornicki P., Scappino L.A., Haselkorn R.;
RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
RT protein.";
RL J. Bacteriol. 175:5268-5272(1993).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L14863; AAA74627.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20496 MW; F3F32A12CF106182 CRC64;

Query Match 53.4%; Score 505; DB 1; Length 185;
Best Local Similarity 52.2%; Pred. No. 3.7e-33;
Matches 96; Conservative 38; Mismatches 50; Indels 0; Gaps 0;

QY 1 MISVNDFTGLTISVDNAIWKVIDFQHVKPGKGSFAVRSKLNRLRTGAIOEKTFRAGEKV 60
DB 1 MISSNDFRPGSVILDGSWRVVIDFLHVKPGKGSFAVRTTLKNVSGKVKLEKTFRAGETV 60

QY 61 EPAMLENRMQVLYADGNHVMFMDNESFEQTELSDDYLKEELNYLKEGMEVQIQTYEGET 120
DB 61 POATLEKITMQHTYKEGDEFVFMDESVEGRLSAAQIGDRVKYLKEGNEVNVIRWGEV 120

QY 121 IGVLPKTVLTVTETEPGIGKGDATGATKSATVTGTYTLNVPFLVNEGDVLIINTGDGS 180
DB 121 LVEVLANSVVLVIQTDPGVKGDATGATGKPAIVETGATVMVPLFISOGRIKIDTRDK 180

QY 181 YISR 184
DB 181 YLGR 184

RESULT 5
EFP_MYCTU
ID EFP_MYCTU STANDARD; PRT; 187 AA.
AC P95019;
DT 01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
ELONGATION FACTOR P (EF-P).
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
SEQUENCE FROM N.A.
STRAIN-H37RV;
MEDLINE=98295987; PubMed=9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
Nature 393:537-544(1998).
-!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
(BY SIMILARITY).
-!- PATHWAY: PROTEIN BIOSYNTHESIS.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; Z83863; CAB06174.1; -.
TUBERCULIST; RV2534C; -.
INTERPRO; IPR001059; -.
PFAM; PF01132; EFP; 1.
PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20407 MW; 58BCBD1DE0424F58 CRC64;

Query Match 50.7%; Score 480; DB 1; Length 187;
Best Local Similarity 47.8%; Pred. No. 3.4e-31;
Matches 88; Conservative 38; Mismatches 58; Indels 0; Gaps 0;

QY 1 MISVNDFTGLTISVDNAIWKVIDFQHVKPGKGSFAVRSKLNRLRTGAIOEKTFRAGEKV 60
DB 1 MATTAADFNGVLVLDGQLWTITEFQHVKPGKGPAPVTRTKLVNLSGKVDKTFNAGVKV 60

QY 61 EPAMLENRMQVLYADGNHVMFMDNESFEQTELSDDYLKEELNYLKEGMEVQIQTYEGET 120
DB 61 DTATVDVRDITYLYRDGSDGFVMDSQDYEQHPLPEALVGDAAARFLLEGMPQVQVAFHNGVP 120

QY 121 IGVLPKTVLTVTETEPGIGKGDATGATKSATVTGTYTLNVPFLVNEGDVLIINTGDGS 180
DB 121 LVIELPVTVELEVTHTEFGLQDRSSAGTKPATLQTGAQINVPFLPINTGDKLVKVDSDGS 180

QY 181 YISR 184
DB 181 YLGR 184

RESULT 6
EFP_BRELA
ID EFP_BRELA STANDARD; PRT; 187 AA.
```

```

AC Q45288;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Brevibacterium lactofermentum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13869;
RX MEDLINE=98036051; PubMed=9370284;
RA Ramos A., Macias J.R., Gil J.A.;
RT "Cloning, sequencing and expression of the gene encoding elongation
RT factor P in the amino-acid producer Brevibacterium lactofermentum
RT (Corynebacterium glutamicum ATCC 13869).";
RL Gene 198:217-222(1997).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99289; CAA67673.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20626 MW; 9E62823094CB0C1B CRC64;

Query Match 48.6%; Score 460; DB 1; Length 187;
Best Local Similarity 48.4%; Pred. No. 1.3e-29;
Matches 89; Conservative 38; Mismatches 57; Indels 0; Gaps 0;

Qy 1 MISVNDKFTGLTISVDNAIKWKVPGKGSFAFVRSKLRNLRTGAIOEKTFRAGEKV 60
   | : ||| | : : : ||| ||| ||| ||| : : : ||| : ||| : |||
Db 1 MATTADEKNGLVKNEGLQIIIEFQHVKPGKGFPAFVRTKLKDVVTGKTIDKTNAGVKV 60

Qy 61 EPAMTENRRMOYLYADGDNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIQTYGETI 120
   | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db 61 ETATVDPRCTLYLNDGTSFIVMDKTFEQVELSPDAGFAGDAGRELLNMRVVOVSFHGEA 120

Qy 121 IGVLPKTVELTVTETEPGKIGDGTATGKTSATVETGYTLNVPFLVNEGDVLIINTGDGS 180
   | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db 121 LFGELPVSVDLRVETDPGLQDGRSTGGTRPATLETGAETQVPLFIETGNVLKVDTRDGS 180

Qy 181 YISR 184
   | : |||
Db 181 YLSR 184

RESULT 7
EFP_AQUAE STANDARD; PRT; 192 AA.
AC O67376;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).

GN EFP OR AQ_1364.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000736; AAC07331.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 192 AA; 21807 MW; 9B651E676507CAA0 CRC64;

Query Match 46.0%; Score 435; DB 1; Length 192;
Best Local Similarity 44.3%; Pred. No. 1.2e-27;
Matches 81; Conservative 38; Mismatches 64; Indels 0; Gaps 0;

Qy 2 ISVNDKFTGLTISVDNAIKWKVPGKGSFAFVRSKLRNLRTGAIOEKTFRAGEKVE 61
   | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db 5 IDINRIKQDIFIEHKGPYRVLDYEHVKGPGKGQAFVRSKLRNLRTGAIOEKTFRAGEKVE 64

Qy 62 PAMTENRRMOYLYADGDNHVFMDNESFEQTELSDDLKEELNLYLKEGMEVQIQTYGETI 121
   | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db 65 LADPEQVYATYSYNDGENYVFMNTQTYDMLAVPKETEEBAKFLKEGMEVIVFLYKGOPI 124

Qy 122 GVLPKTVELTVTETEPGKIGDGTATGKTSATVETGYTLNVPFLVNEGDVLIINTGDGSY 181
   | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : |||
Db 125 GIELPKHVELOVWETEPAFKGDQAGTKPAKLETGAVIQVPPFVKEGDVKVDTRTGSY 184

Qy 182 ISR 184
   | : |||
Db 185 VER 187

RESULT 8
EFP_THEMA STANDARD; PRT; 185 AA.
AC Q9X284;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP OR TM1763.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=MSB8 / DSM 3109;
RA MEDLINE=95287316; PubMed=10360571;
RX Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001815; AAD36827.1; -
DR TIGR; TM1763; -
DR PROSITE; PS01275; EPP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20875 MW; A18A98DF3597A68E CRC64;

Query Match 44.5%; Score 421; DB 1; Length 185;
Best Local Similarity 44.6%; Pred. No. 1.5e-26;
Matches 82; Conservative 35; Mismatches 67; Indels 0; Gaps 0;

Qy 1 MSVNDKFTGLTISVDNAIKWVIDFQHVKPGKSAFVRSKLRNLTGAIQKTRFAGEKV 60
Db 1 MIEVGDLKKGFLIYDGEIYRVLEASKHFMGSGSLIRTKLNKVRTPFVREVNFPSEGV 60
Qy 61 EPAMTENRMQVLYADGNHVMFMDNESFEQTELSDDLKEELNLYLKEGMEVQIQTYEGET 120
Db 61 QEAELSFRKAQVLYRDGDHYFMTLDDYEQVASEEIGDAKYILVENKVDLVFHEGTP 120
Qy 121 IGVELPKTVELTVTEPGIKGDTATGATKSAVTGTGTLNVLPLFVNEGDVLLINTGDSG 180
Db 121 IGIELPTVELTVTEPSEFSGKDTVSGGKPAVLETGLKITVPIFYIEVGDRIKVDTRTGE 180
Qy 181 YISR 184
Db 181 YVGR 184

RESULT 9
EPP_HELPY
ID EPP_HELPY STANDARD; PRT; 187 AA.
AC P56004;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EPP OR HP0177.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;

```

```

RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Cocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000538; AAD07247.1; -
DR TIGR; HP0177; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EPP; 1.
DR PROSITE; PS01275; EPP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 187 AA; 20788 MW; 1B433EDF0FAB32B CRC64;

Query Match 43.2%; Score 409; DB 1; Length 187;
Best Local Similarity 40.4%; Pred. No. 1.3e-25;
Matches 74; Conservative 45; Mismatches 64; Indels 0; Gaps 0;

Qy 2 ISVNDKFTGLTISVDNAIKWVIDFQHVKPGKSAFVRSKLRNLTGAIQKTRFAGEKVE 61
Db 3 IGMSELKGLKIELGGVPYRIVEYQHVKPGKGAAPVRAKIKSFLDGKVIKTFHAGDKCE 62
Qy 62 PAMTENRMQVLYADGNHVMFMDNESFEQTELSDDLKEELNLYLKEGMEVQIQTYEGETI 121
Db 63 EPNLVKTMQVLYHDGDTYQFMDTESYEQIALNDSQVGEASKWMLDGMQVQLLNDRKAI 122
Qy 122 GVLEPKTVELTVTEPGIKGDTATGATKSAVTGTGTLNVLPLFVNEGDVLLINTGDSY 181
Db 123 SVDVPQVVALKIVETAPNFKGDTSSASKKPALETGAVVQVFPFVLEGEIILKVNTEY 182
Qy 182 ISR 184
Db 183 LEK 185

RESULT 10
EPP_BACFR
ID EPP_BACFR STANDARD; PRT; 185 AA.
AC P70889;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EPP.
OS Bacteroides fragilis.
OC Bacteria; CFB group; Bacteroidaceae; Bacteroides.
RN [1]
RP SEQUENCE FROM N.A.

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 14:58:00 ; Search time 39.67 Seconds
(without alignments)
546.596 Million cell updates/sec

Title: US-09-322-732-1
Perfect score: 946
Sequence: 1 MISVNDKFTGLTISVDNAIW.....VNEGDVLIINTGSGYSIRG 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_15.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organellae.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	730	77.2	185	2 Q9K951	Q9K951 bacillus ha
2	487.5	51.5	188	2 Q9KX09	Q9KX09 streptomyc
3	443	46.8	185	2 Q9RV32	Q9RV32 deinococcus
4	439	46.4	184	10 Q9M917	Q9M917 arabidopsis
5	434	45.9	189	2 Q9PHW3	Q9PHW3 campylobact
6	379.5	40.1	190	2 Q9PLH1	Q9PLH1 chlamydia m
7	372	39.3	193	2 Q9PAM3	Q9PAM3 xylella fas
8	371	39.2	188	2 Q9KNS1	Q9KNS1 vibrio chol
9	363.5	38.4	190	2 Q9J0B4	Q9J0B4 chlamydia p
10	343.5	36.3	188	2 Q9PQJ3	Q9PQJ3 ureaplasma
11	332	35.1	136	2 Q9S3F5	Q9S3F5 helicobacte
12	330	34.9	136	2 Q9S3F9	Q9S3F9 helicobacte
13	330	34.9	136	2 Q9R3F4	Q9R3F4 helicobacte
14	330	34.9	136	2 Q9R2M0	Q9R2M0 helicobacte
15	328	34.7	187	2 Q9RNJ9	Q9RNJ9 zymomonas m
16	326	34.5	136	2 Q9S3F8	Q9S3F8 helicobacte
17	326	34.5	136	2 Q9S3F7	Q9S3F7 helicobacte
18	326	34.5	136	2 Q9S3F6	Q9S3F6 helicobacte
19	324	34.2	161	2 Q69281	Q69281 citrobacter

20	321	33.9	136	2 Q9S3G0	Q9S3G0 helicobacte
21	291	30.8	185	2 Q9PKR6	Q9PKR6 chlamydia m
22	273.5	28.9	199	2 Q9KSP7	Q9KSP7 vibrio chol
23	270	28.5	185	2 Q9JQ76	Q9JQ76 chlamydia p
24	266.5	28.2	186	2 Q9JUJ2	Q9JUJ2 neisseria m
25	265.5	28.1	186	2 Q9JZQ8	Q9JZQ8 neisseria m
26	225.5	23.8	189	2 Q9PBE1	Q9PBE1 xylella fas
27	225	23.8	109	2 P82459	P82459 streptococc
28	189.5	20.0	212	10 Q9STQ4	Q9STQ4 arabidopsis
29	99.5	10.5	138	1 Q9VOM2	Q9VOM2 pyrococcus
30	95	10.0	731	5 Q9W148	Q9W148 drosophila
31	94.5	10.0	377	2 Q9K6U6	Q9K6U6 bacillus ha
32	89	9.4	277	2 Q9XDY4	Q9XDY4 unidentified
33	88	9.3	753	5 Q9VLZ0	Q9VLZ0 drosophila
34	88	9.3	798	5 Q9VLY9	Q9VLY9 drosophila
35	88	9.3	1420	5 Q9NEV7	Q9NEV7 drosophila
36	87.5	9.2	1150	2 Q9K9M0	Q9K9M0 bacillus ha
37	86.5	9.1	142	10 Q93368	Q93368 brassica ol
38	86.5	9.1	142	10 Q41213	Q41213 brassica na
39	86.5	9.1	550	5 Q9VFW3	Q9VFW3 drosophila
40	86	9.1	2367	2 Q46034	Q46034 clostridium
41	85.5	9.0	458	2 P71472	P71472 lactobacill
42	85.5	9.0	546	12 Q96682	Q96682 simian cyto
43	85.3	9.0	823	2 Q9KE85	Q9KE85 bacillus ha
44	85	9.0	680	2 Q52644	Q52644 ruminococcu
45	84.5	8.9	785	10 Q64756	Q64756 arabidopsis

ALIGNMENTS

RESULT 1	
ID Q9K951	PRELIMINARY; PRT; 185 AA.
AC Q9K951	Q1-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000	(TREMREL. 15, Last sequence update)
DT 01-OCT-2000	(TREMREL. 15, Last annotation update)
DE TRANSLATION	ELONGATION FACTOR EF-P.
GN EFP	
OS Bacillus halodurans.	
OC Bacteria; Firmicutes; Bacillus/Clostridium group;	
OC Bacillus/Staphylococcus group; Bacillus.	
OX NCBI_TaxID=86665;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN-C-125 / JCM 9153;	
RC Takami H., Nakasone K., Takaki Y.;	
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AP001516; BAB06518.1; -.	
KW Elongation factor.	
SQ SEQUENCE 185 AA; 20566 MW; 8CC48498BD6B37E3 CRC64;	

Query Match	77.2%; Score 730; DB 2; Length 185;
Best Local Similarity	75.0%; Pred. No. 2e-50;
Matches 138; Conservatively	21; Mismatches 25; Indels 0; Gaps 0;
QY 1 MISVNDKFTGLTISVDNAIWVIDFQHVKPGKSAFVRSKLRNLRPTGAIQKTFRAGEKV 60	
Db 1 MISVNDKFTGLTIEVDNGIQVMEFQHVKPGKGAFAVRSKLRNLRPTGAVQKTFRAGEKV 60	
QY 61 EPAMIEHNRMOYLYADGDNHVFMDNESFEQTELSVDYLKKEELNLYLKEGMEVQIOFYEGET 120	
Db 61 SKAIEHNRMOYLYASGDVHFTFMDNETFEQLELSTAQIEHLEKFKENMEVHVHVSIOGET 120	
QY 121 IGVLPKPTVELTVTETEPGKNGDPTATGATKTSATVETGYTLNVPFLFVNEGDVLIINTGDGS 180	
Db 121 LGVEVPNTVELTVTETEPGKNGDPTASGTRKPTLETGLTVQVPPFVNEGDVLIIDTRSGD 180	
QY 181 YISR 184	
Db 181 YVSR 184	

```

RESULT 2
ID Q9KXQ9 PRELIMINARY; PRT; 188 AA.
AC Q9KXQ9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN EFP.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RL "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL357523; CAB93371.1; -.
SQ SEQUENCE 188 AA; 20649 MW; 3058EFB64B688C1C CRC64;

Query Match 51.5%; Score 487.5; DB 2; Length 188;
Best Local Similarity 48.1%; Pred. No. 2.9e-31;
Matches 89; Conservative 38; Mismatches 57; Indels 1; Gaps 1;

QY 1 MISVNDFKTGLTISVDNA-IWKVIDFOHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEK 59
Db 1 MASTNDLKNLKGQLGGLQSVVVFQVRKPGKPAFVRTKLNVLGSKVYVDKTFNAGVK 60

QY 60 VEPAMENRRMOYLYADGDHNVFMDNESFEQTELSDDLKEELNLYLKEGMEVOIQTYE 119
Db 1 VETATVDKRMQFSYMDGDYFVMDMETYDQLMIDRKVVGDAANFLVEGFATVAQHEGE 120

QY 120 TIGVELPKTVLTETEPGKIGDTATGATKSATVETGYTLNPLFVNEGDLVLIINTGDG 179
Db 121 VLFVELPAAVELTIQETEPGVQGRDSTGGTKPATLETGHQINVPFLFITTEKIKVDTRTS 180

QY 180 SYISR 184
Db 181 DYLR 185

RESULT 3
Q9RY32 PRELIMINARY; PRT; 185 AA.
AC Q9RY32;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P.
GN DR0119.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=RL;

```

```

RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamohile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Dali M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans RL.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001874; AAF09709.1; -.
DR TIGR; DR0119; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Elongation factor.
SQ SEQUENCE 185 AA; 20475 MW; 893FD0850AC102D3 CRC64;

Query Match 46.8%; Score 443; DB 2; Length 185;
Best Local Similarity 40.8%; Pred. No. 9.3e-28;
Matches 75; Conservative 49; Mismatches 60; Indels 0; Gaps 0;

QY 1 MISVNDFKTGLTISVDNAIWKVIDFOHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEKV 60
Db 1 MISVTELRNGTKVQMDGLWECLDYSHLKMGRGKAVVTKFRNMESGSIVDRTFNSTEKL 60

QY 61 EPAMENRRMOYLYADGDHNVFMDNESFEQTELSDDLKEELNLYLKEGMEVOIQTYE 120
Db 61 QDTYVEGKMKOYLYPDGDYFVMDMETFDQVHLGKNIIVDAAKFMKENTEEVAMYGDKA 120

QY 121 IGVELPKTVLTETEPGKIGDTATGATKSATVETGYTLNPLFVNEGDLVLIINTGDG 180
Db 121 LSISLPNQVILKITQDTPGVGRDTSVGGTKPATLETGAVVQVPLFVEQGDVVDRTGQ 180

QY 181 YISR 184
Db 181 YLSR 184

RESULT 4
Q9M9I7 PRELIMINARY; PRT; 184 AA.
AC Q9M9I7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE ELONGATION FACTOR P (EF-P).
GN F17014.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F17014 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012562; AAF26065.1; -.
KW Elongation factor.
SQ SEQUENCE 184 AA; 20599 MW; CFC2A79C3485BAAD CRC64;

Query Match 46.4%; Score 439; DB 10; Length 184;
Best Local Similarity 45.4%; Pred. No. 1.9e-27;
Matches 83; Conservative 39; Mismatches 61; Indels 0; Gaps 0;

QY 2 ISVNDFKTGLTISVDNAIWKVIDFOHVKPGKGSFVRSKLRNLRTGAIQKTFRAGEKV 61

```



```

Db 1 MSANDIKAGTNIIEVDGAPWRVLEFLHVPKGGAAAFVTKIRNYVNGSTVETFRAGISVE 60
QY 62 PAMIEENRRMOYLYADGDNHVFMDSFOTELSSDYLKKEELNYLKEGMEVOIQTEGETI 121
Db 61 EANIYKETQFYTKDGSQFVMDLTYETRLNESDMGEKTKWLKEGMDCLILYWKDKVI 120
QY 122 GVLPKTVLVTETEPGKIGDTATGATKATSVETGYTLNVLPLFVNEGVDVLIINTGDSY 181
Db 121 DFDLPIITVKLKVDVDPGLRGDTVGQSKPATMGTGAIVAVPLFINVGEIEFVDTRTCAY 180
QY 182 ISR 184
Db 181 MNR 183

RESULT 5
Q9PHW3 PRELIMINARY; PRT; 189 AA.
ID Q9PHW3
AC Q9PHW3;
DT 01-OCT-2000 (TremBrel. 15, Created)
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)
DE 01-OCT-2000 (TremBrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN EFP.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group.
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holt R.D.,
RA Jorgensen K., Kiley M., Kiley M., Kiley M., Kiley M., Kiley M.,
RA Quail M.A., Rajandream A.M., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.
RL Nature 403:665-668(2000).
DR EMBL: AL139075; CAB75187.1; -.
DR INTERPRO: IPR001059; -.
DR PROSITE: PS01184; -.
DR PFAM: PF01132; EFP; 1.
DR PFAM: PF01287; eIF-5a; 1.
SQ SEQUENCE 189 AA; 21098 MW; C0BAE2CC7D220300 CRC64;

Query Match 45.9%; Score 434; DB 2; Length 189;
Best Local Similarity 45.1%; Pred. No. 4.9e-27;
Matches 82; Conservative 37; Mismatches 63; Indels 0; Gaps 0;

QY 3 SVNDFKTGLTISVDNAIKWIDFQHVKPGKGSFVRSKLNRLRTGAIQKTFRAGEKVEP 62
Db 5 SMGDLKGLKIDIGIPKEIVQHVKPGKGFVRIKIKSFIDKGLKTFHAGDKCEA 64
QY 63 AMIENRRMOYLYADGDNHVFMDSFOTELSSDYLKKEELNYLKEGMEVOIQTEGETI 122
Db 65 PNLDEKTMQYLDGGENQFMDQTYEQVAISDDVDGAKKWLMDGMVDFLHNGKAIG 124
QY 123 VELPKTVLVTETEPGKIGDTATGATKATSVETGYTLNVLPLFVNEGVDVLIINTGDSY 182
Db 125 VEPQVVELKIETAPNFKGDTQGSNKPATLETCAVQIPHPVLEGEVIRVDTVRGEYI 184
QY 183 SR 184
Db 185 ER 186

RESULT 6
Q9PLH1 PRELIMINARY; PRT; 190 AA.
ID Q9PLH1
```

```

AC Q9PLH1;
DT 01-OCT-2000 (TremBrel. 15, Created)
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)
DE 01-OCT-2000 (TremBrel. 15, Last annotation update)
DE TRANSLATION ELONGATION FACTOR P.
GN TC0133.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
RA Gwinn K., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Read T.D., Brunham R., Shen C., Gill S.R., Heidelberg J.F., White O.,
RA Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn K., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE002280; AAF39011.1; -.
DR TIGR: TC0133; -.
DR INTERPRO: IPR001059; -.
DR PFAM: PF01132; EFP; 1.
DR PROSITE: PS01275; EFP; 1.
SQ SEQUENCE 190 AA; 21550 MW; C5920D53944CD9BD CRC64;

Query Match 40.1%; Score 379.5; DB 2; Length 190;
Best Local Similarity 40.8%; Pred. No. 9.9e-23;
Matches 75; Conservative 40; Mismatches 68; Indels 1; Gaps 1;

QY 2 ISVNDFKTGLTISVDNAIKWIDFQHVKPGKGSFVRSKLNRLRTGAIQKTFRAGEKVE 61
Db 4 VSTSEFRVRLGVEIDGQPYVILQNDVFKPGKGFVRIKIKSFIDKGLKTFHAGDKCEA 63
QY 62 PAMIEENRRMOYLYADGDNHVFMDSFOTELSSDYLKKEELNYLKEGMEVOIQTEGETI 121
Db 64 TADVREQMRLLYTDOEGATFMDDETPEQELIFWDKLENIQWLEDTVYTLVRYNGDVI 123
QY 122 GVLPKTVLVTETEPGKIGDTATG-ATKATSVETGYTLNVLPLFVNEGVDVLIINTGDS 180
Db 124 SVPEPIFNELSIAETAPGVGDTASGRVLPKATNTGAKINVPFIIEGSEVVKVDTRTGS 183
QY 181 YISR 184
Db 184 YESR 187

RESULT 7
Q9PAM3 PRELIMINARY; PRT; 193 AA.
ID Q9PAM3
AC Q9PAM3;
DT 01-OCT-2000 (TremBrel. 15, Created)
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)
DE 01-OCT-2000 (TremBrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN XF2473.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kurane E.E., Laigret F., Lambais M.R., Leite L.C.P.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miraca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Ouaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zaco M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
RL Nature 406:151-157(2000).
DR EMBL; AE004055; AAF85271.1; -;
DR INTERPRO; IPR001059; -;
DR PFAM; PF01132; EFP; 1.
SQ SEQUENCE 193 AA; 21380 MW; 8D251D1472A37D06 CRC64;

Query Match 39.3%; Score 372; DB 2; Length 193;
Best Local Similarity 40.9%; Pred. No. 4e-22;
Matches 74; Conservative 40; Mismatches 67; Indels 0; Gaps 0;
Qy 4 VNDFKTLTISVDNAIKWIDFQHVKGKGSFAFVRSKLNRLRTGAIQKTPRAGEKVEPA 63
Db 11 MNDVKNKMLVNAEPAVITDTEVYKPGKQAFTRVYRLKSRGVQEVTKMKSTDILEAA 70
Qy 64 MIENRRMOYLXADGNHVMFNDSFEQTELSDDLKEELNYLKEGMEVQIOTYEGETIGV 123
Db 71 DVVDTDQYLYSDGEYWHFMQETFEQVQADKNGMGAEKWLKEEQCVVTLWNGVPIGV 130
Qy 124 ELPKTVLTVTETPGIKGDTATGATKSATVETGYTLNVPFLVNEGDVLIINTGDGYSIS 183
Db 131 QPPNPFVKLTETDPLRGDTSGGGKGPATLETGAVRVPLFVQDEVIKVDTRSGEYVS 190
Qy 184 R 184
Db 191 R 191
RESULT 8
Q9KNS1 PRELIMINARY; PRT; 188 AA.
AC Q9KNS1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ELONGATION FACTOR P.
GN VC2660.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smit H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.";
RL Nature 406:477-483(2000).
DR EMBL; AE004332; AAF95801.1; -;
DR TIGR; VC2660; -;
SQ SEQUENCE 188 AA; 20576 MW; 73537EA64626B9A2 CRC64;
Query Match 39.2%; Score 371; DB 2; Length 188;
Best Local Similarity 41.0%; Pred. No. 4.6e-22;
Matches 75; Conservative 35; Mismatches 72; Indels 0; Gaps 0;
Qy 2 ISVDFKTLTISVDNAIKWIDFQHVKGKGSFAFVRSKLNRLRTGAIQKTPRAGEKVE 61
Db 4 VSTNEFKGLKIMLDNEPCVILENEYKPGKQAFNRVIRKLLTGKVLKTFKSGDTAE 63
Qy 62 PAMTENRRMOYLXADGNHVMFNDSFEQTELSDDLKEELNYLKEGMEVQIOTYEGETI 121
Db 64 VADVVDIDLYNDGFEYHFMNNSFEQLAADAQAVGENAKWLVENNTCMLTLWNGNPI 123
Qy 122 GVLPKTVLTVTETPGIKGDTATGATKSATVETGYTLNVPFLVNEGDVLIINTGDGYS 181
Db 124 AVTPNPFVELEVETETDPGVKGDGTGGKGPATLSTGAVRVPLFVQIGEVIKVDTRSAEY 183
Qy 182 ISR 184
Db 184 VGR 186

RESULT 9
Q9JQB4 PRELIMINARY; PRT; 190 AA.
AC Q9JQB4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ELONGATION FACTOR P (TRANSLATION ELONGATION FACTOR P).
GN EFP_2 OR CP0971.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83538;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of *Chlamydia pneumoniae* J138 from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,
RA Barry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwinn M., Nelson W., Deboy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AP002548; BAA99103.1; -;
DR TIGR; CP0971; -;
KW Elongation factor.

RESULT
Q9S3F5
11

```
Query Match          34.9%; Score 330; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 5.1e-19;
Matches 61; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

QY 25 FOHVKPGKGSFAVRKSLRNLTGAIQKTFRAGEKVEPAMENRRMOYLYADGDNHVFMD 84
    :|||||:|||||: : : ||| | : : : ||| | : : : ||| | : : : ||| |
Db 1 YQHVKPGKGAFAVRKISFLDKGVIEKTFHAGDKCEPNLVKTMQYLYHGDGTQYFMD 60

QY 85 NESFEQTELSSDYKLKEELNYLKEGMEVQIOTYEGETIGVLPKTVLTETETEPGKGD 144
    ||:| | | : : : ||:| | | : : : ||:| | | : : : ||:| | |
Db 61 IESYEQIALNDSQVGEASKWMLDGMQVLLHNDKAISVDVPQVVALKIVETAPNFKGDT 120

QY 145 ATGATKSATVETG 157
    : : : | ||:| |
Db 121 SSASKKPATLETG 133

RESULT 13
Q9R3F4 PRELIMINARY; PRT; 136 AA.
AC Q9R3F4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5-1, AND NCTC11638;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239623; CAB37734.1; -
DR EMBL; AJ239610; CAB37721.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
DR Elongation factor.
KW NON_TER 1
FT NON_TER 136
FT NON_TER 136
SQ SEQUENCE 136 AA; 7E4970D8D2896C92 CRC64;
```

```
Query Match          34.9%; Score 330; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 5.1e-19;
Matches 61; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

QY 25 FOHVKPGKGSFAVRKSLRNLTGAIQKTFRAGEKVEPAMENRRMOYLYADGDNHVFMD 84
    :|||||:|||||: : : ||| | : : : ||| | : : : ||| | : : : ||| |
Db 1 YQHVKPGKGAFAVRKISFLDKGVIEKTFHAGDKCEPNLVKTMQYLYHGDGTQYFMD 60

QY 85 NESFEQTELSSDYKLKEELNYLKEGMEVQIOTYEGETIGVLPKTVLTETETEPGKGD 144
    ||:| | | : : : ||:| | | : : : ||:| | | : : : ||:| | |
Db 61 IESYEQIALNDSQVGEASKWMLDGMQVLLHNDKAISVDVPQVVALKIVETAPNFKGDT 120

QY 145 ATGATKSATVETG 157
    : : : | ||:| |
Db 121 SSASKKPATLETG 133

RESULT 14
Q9R2M0 PRELIMINARY; PRT; 136 AA.
ID Q9R2M0;
AC Q9R2M0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
```

```
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239625; CAB37736.1; -
DR EMBL; AJ239613; CAB37724.1; -
DR EMBL; AJ239614; CAB37725.1; -
DR EMBL; AJ239615; CAB37726.1; -
DR EMBL; AJ239616; CAB37727.1; -
DR EMBL; AJ239617; CAB37728.1; -
DR EMBL; AJ239618; CAB37729.1; -
DR EMBL; AJ239619; CAB37730.1; -
DR EMBL; AJ239620; CAB37731.1; -
DR EMBL; AJ239621; CAB37732.1; -
DR EMBL; AJ239622; CAB37733.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
DR Elongation factor.
KW NON_TER 1
FT NON_TER 136
FT NON_TER 136
SQ SEQUENCE 136 AA; 15019 MW; FCC35872F287829C CRC64;

Query Match          34.9%; Score 330; DB 2; Length 136;
Best Local Similarity 45.9%; Pred. No. 5.1e-19;
Matches 61; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

QY 25 FOHVKPGKGSFAVRKSLRNLTGAIQKTFRAGEKVEPAMENRRMOYLYADGDNHVFMD 84
    :|||||:|||||: : : ||| | : : : ||| | : : : ||| | : : : ||| |
Db 1 YQHVKPGKGAFAVRKISFLDKGVIEKTFHAGDKCEPNLVKTMQYLYHGDGTQYFMD 60

QY 85 NESFEQTELSSDYKLKEELNYLKEGMEVQIOTYEGETIGVLPKTVLTETETEPGKGD 144
    ||:| | | : : : ||:| | | : : : ||:| | | : : : ||:| | |
Db 61 IESYEQIALNDSQVGEASKWMLDGMQVLLHNDKAISVDVPQVVALKIVETAPNFKGDT 120

QY 145 ATGATKSATVETG 157
    : : : | ||:| |
Db 121 SSASKKPATLETG 133

RESULT 15
Q9RNJ9 PRELIMINARY; PRT; 187 AA.
ID Q9RNJ9;
AC Q9RNJ9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P EFP.
GN EFP.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
OC Zymomonas.
OX NCBI_TaxID=542;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZM4;
RA Lee H.J., Kang H.S.;
RT "Sequence analysis of 42P4 fosmid clone of Zymomonas mobilis ZM4.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF180145; AAD56928.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
```

Search completed: January 12, 2001, 15:00:08
Job time: 128 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: January 12, 2001, 14:58:32 ; Search time 27.86 Seconds
(without alignments)
230.741 Million cell updates/sec
Title: US-09-322-732-2
Perfect score: 975
Sequence: 1 MATYNSNDFRAGLKIMLDGE.....QIGEVIKVTRSEYVSRVK 188
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36: *
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Query		DB		Description	
Result No.	Score	Match	Length	ID					
1	434.5	46.6	190	20	Y37395			Protein involved i	
2	445.5	45.7	190	20	Y35550			Chlamydia pneumoni	
3	362	37.1	185	21	Y70152			Staphylococcus aur	
4	229	23.5	185	20	Y37084			Protein involved i	
5	226	23.2	185	20	Y34781			Chlamydia pneumoni	
6	78	8.0	605	20	Y16791			Truncated plant-op	
7	78	8.0	718	19	W61024			Chimeric lepidopte	
8	78	8.0	1148	16	R84731			CryIF/cryIA(b) chi	
9	78	8.0	1148	16	R84733			CryIF/cryIA(b) chi	
10	78	8.0	1148	17	R94907			CryIF/cryIA(b) chi	
11	78	8.0	1148	19	W76718			Alternative cryIF/	
12	78	8.0	1148	19	W76717			Alternative cryIF/	

13	78	8.0	1148	19	W76707	Plasmid pMYC2244 c
14	78	8.0	1148	19	W76709	Plasmid pMYC2523 c
15	78	8.0	1148	19	W76715	Consensus cryIF/cr
16	78	8.0	1148	20	Y16790	Plant-optimised cr
17	78	8.0	1150	16	R84734	CryIA(c)/cryIF/cry
18	78	8.0	1150	17	R94913	CryIA(c)/cryIF/cry
19	78	8.0	1150	19	W76706	Plasmid pMYC2239 b
20	78	8.0	1174	12	R10131	Lepidopteran-activ
21	78	8.0	1174	12	R14856	CryIF protein. Ba
22	78	8.0	1174	14	R39754	Delta endotoxin.
23	78	8.0	1174	16	R84735	CryIF toxin with l
24	78	8.0	1174	16	R84732	CryIF/436 chimeric
25	78	8.0	1174	17	R94908	CryIF/436 chimeric
26	78	8.0	1174	17	R94914	CryIF toxin encode
27	78	8.0	1174	17	R94916	Native cryIF/CryIA
28	78	8.0	1174	19	W76708	Plasmid pMYC2243 B
29	78	8.0	1174	19	W76710	Plasmid pMYC2254 c
30	78	8.0	1174	19	W76712	B. thuringiensis c
31	78	8.0	1174	20	Y16792	Wild-type B.t cryI
32	78	8.0	1175	19	W76716	Alternative cryIF/
33	78	8.0	1184	16	R84737	Native cryIF toxin
34	77	7.9	609	21	Y58824	E. coli glucosamin
35	76	7.8	609	21	Y58822	E. coli glucosamin
36	76	7.8	609	21	Y58823	E. coli glucosamin
37	76	7.8	609	21	Y58825	E. coli glucosamin
38	76	7.8	609	21	Y58826	E. coli glucosamin
39	76	7.8	609	21	Y58827	E. coli glucosamin
40	75	7.7	170	21	Y59205	Caldicellulosirupt
41	74	7.6	613	21	Y92512	Human OXRE-9. Ho
42	73.5	7.5	1098	17	R99392	Haemophilus adhesi
43	73	7.5	748	21	Y70231	Human RNA-associat
44	73	7.5	1215	21	Y55791	Superheat-resistan
45	72.5	7.4	299	18	W55450	H. pylori ORF 02ae

ALIGNMENTS

RESULT 1	
Y37395	
ID Y37395 standard; Protein; 190 AA.	
XX	
AC Y37395;	
XX	
DT 07-OCr-1999 (first entry)	
XX	
DE Protein involved in transcription, translation and/or maturation.	
XX	
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;	
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;	
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;	
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.	
XX	
OS Chlamydia trachomatis.	
XX	
PN WO9928475-A2.	
XX	
PD 10-JUN-1999.	
XX	
PF 27-NOV-1998; 98WO-IB01939.	
XX	
PR 04-NOV-1998; 98US-0107077.	
PR 28-NOV-1997; 97FR-0015041.	
PR 17-DEC-1997; 97FR-0016034.	
XX	
PA (GEST) GENSET.	
XX	
PI Griffais R;	
XX	
DR WPI; 1999-371125/31.	
XX	
PT Genome sequence of Chlamydia trachomatis	
XX	

CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the *C. pneumoniae* genome (see Y34584-Y35879) can be used
CC immunogenic compositions as vaccines. Vectors containing *C. pneumoniae*
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of *C. pneumoniae*.
XX Sequence 185 AA;
SQ

Query Match	23.2%;	Score	226;	DB	20;	Length	185;
Best Local Similarity	32.8%;	Pred. No.	2.6e-17;				
Matches	60;	Conservative	33;	Mismatches	87;	Indels	4; Gaps
2;							
QY	6	SNDFRAGLKTIMLDCEPYAEASBFVKPGKQAARVAKLRLLTGTTRVEKTEFKSTDSABGA	65				
Db		: : : : : : : : : : : : : : : : : : : :					
QY	4	ssqlsvgmfiistkdglvklvsakvaqgesfikvalqaadsdvviernfkatcqvekea	63				
Db		: : : : : : : : : : : : : : : : : : : :					
QY	66	DVYDMNLTYLYNDGEFWHFNNETFEOLSDADAKAIGNAKWLLDQACICTVLNNGQPISV	125				
Db		: : : : : : : : : : : : : : : : : : : :					
QY	64	qfetrtleylyledesylfdlgnyeklifpqeimkdnflfkagvtvsvamvydnnvfsv	123				
Db		: : : : : : : : : : : : : : : : : : : :					
QY	126	TPPNFVELETVTDPLGLKGDTAGTGG--KPATLSTGAAVVPLEPVGQIEVTIKVPTSRGEY	183				
Db		: : : : : : : : : : : : : : : : : : : :					
QY	124	elphfielmwsktd--fpqdslslsggvkkalletgtievmpvpfieigdvikiditrccey	181				
Db		: : : : : : : : : : : : : : : : : : : :					
QY	184	VSRV	187				
Db		: :					
QY	182	iqr	185				
Db		:					

RESULT	6
ID	Y16791
XX	Y16791 standard; Protein; 605 AA.
XX	AC
XX	AC
XX	Y16791;
DT	27-JUL-1999 (first entry)
DT	
XX	
XX	Truncated plant-optimised cryIF toxin.
DE	
XX	Plant-optimised gene; pesticial toxin; CryIF; CryIAC; CryIC; I58C2c;
KW	3lG1a; CryIAB protoxin; insect control; chimeric.
KW	
XX	
XX	Bacillus thuringiensis.
OS	Synthetic.
OS	
XX	
PN	WO9924581-A2.
PN	
XX	
XX	20-MAY-1999.
XX	
XX	
PF	04-NOV-1998; 98WO-US23457.
XX	
XX	23-OCT-1998; 98US-0178252.
PR	12-NOV-1997; 97US-0065215.
PR	02-MAR-1998; 98US-0076445.
XX	
XX	
PA	(MYCO) MYCOGEN CORP.
XX	
PI	Cardineau GA, Narva KE, Stelman SJ;

xx Nucleic acids encoding pesticidal proteins, optimized for plant
 pt expression
 xx Claim 7; Page 27-29; 67pp; English.
 PS

CCC The invention relates to nucleic acid sequences, optimised for expression in plants, that encode: (i) one of the pesticidal toxins CryIIA, CryIAC, CryIIIC

CryIc, 158C2c or 31G1a, or (ii) a portion of the CryIAB protoxin. The pesticial toxins (Y16790, Y16792-Y16798) are used for pest, especially insect control, particularly on plants. The nucleic acid sequences are used to produce (truncated or chimeric) toxins, and related fusion genes or proteins, particularly they are used to generate transgenic plants that show increased resistance to pests. Compared with wild-type Bacillus thuringiensis sequences, the optimised sequences are better suited to expression in plants. The present sequence represents an amino acid sequence for a truncated, plant optimised cryIF toxin.

Sequence 605 AA;
SQ

	Query Match	8.0%;	Score 78;	DB 20;	Length 605;
	Best Local Similarity	23.4%;	Pred. No. 3.6;		
	Matches 39;	Conservative 29;	Mismatches 67;	Indels 32;	Gaps 9;
QY	18	DGEPYAVEASE--FVAPGKGQAFARVKLRLL---TGTVERKETKSDSAEGADVDDMNL	72		
Db	351	dprpfyrtlsdpfvrggfgonphvylgrvafqgtgnhtrtrfnsqtdlsdeipqp-	409		
QY	73	TYLYNDGEFW----HPMNNFTFQLSADAKAIGDNKAWLLDQABCIYTLNMGQPSIVTTP	128		
Db	410	---dngsapwndyshlvnhvtfvrwpg---isgdsdw-----rapmfswth--satpt	456		
QY	129	NFVELEIVDTDPLKGDTAGTGKGPATLSTGANVVKVPLFVQIGEVIK	175		
Db	457	ntidperitqilpvkahlqsg-----ttvrrpgftg-gailr	494		

	7	RESULT
W61024		
ID	W61024 standard; protein; 718 AA.	
XX		
AC	W61024;	
XX		
DT	23-SEP-1998 (first entry)	
XX		
DE	Chimeric lepidopteran-toxic crystal protein CryIF core toxin portion.	
XX		
KW	Chimeric; lepidopteran-toxic crystal protein; CryIF; CryIac; insect;	
KW	core toxin; tail toxin; <i>Bacillus thuringiensis</i> ; insecticide.	

OS Synthetic.
OS Chimeric - *Bacillus thuringiensis*.

XX
PN
WO9815170-A1.

XX
PD 16-APR-1998.

XX
DE
08-OCT-1997:

XX	00	001	1997,	57	00M/	0019100
XX	00	003	1997	00	00M/	0019100

PR 08-OCT-1996; 960S-0/31079.
XX

PA (ECOG-) ECOGEN INC.
XX

PI Baum JA, Gilmer AJ;
VY

WPI; 1998-322275/28.

Chimeric CryIF-CryIA

PT insect pests
XX

PS Claim 7; Page 33-35; 55pp; English.
yy

The present sequence represents a chimeric protein comprising a CryIF core toxin-containing portion. Chimeric proteins have been developed comprising a CryIF core toxin-containing portion and a CryIAC tail-containing portion. The present invention also describes: (1) an isolated *Bacillus thuringiensis* (B.t.) producing a chimeric protein as defined above, where the *Bacillus* has normal sporulation; (2) a lepidopteran toxic crystal protein produced by the microorganism of (1); (3) a microorganism spore comprising a nucleic acid (NA) encoding a

DR	N-PSDB; T05249.
XX	
XX	Nucleic acid encoding chimeric Bacillus thuringiensis
PT	delta-endotoxin - providing increased expression in Pseudomonas,
PT	esp. for control of lepidoptera pests.
XX	
XX	Claim 6; Page 42-45; 9lpp; English.
XX	
CC	The sequence represents the cryIF/cryIA(b) chimeric toxin encoded
CC	by plasmid pMYC2244. The transition from cryIF core toxin
CC	portion to heterologous protoxin portion occurs after the
CC	characteristic CryI toxin sequence of R84729 and before the end
CC	of peptide R84730. The chimeric toxin is expressed in Pseudomonas
CC	fluorescens better than native delta endotoxins. Host cells
CC	expressing the chimeric gene and producing chimeric toxin may be
CC	used in insecticide compositions. Where the host cells are plant
CC	cells, the gene confers insect resistance to the transformed
CC	plant.
XX	
XX	Sequence 1148 AA;
SQ	
	Query Match 8.0%; Score 78; DB 16; Length 1148;
	Best Local Similarity 23.4%; Pred.No. 9.1;
	Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps
QY	18 DGPYAVEASE--FVKPRGKGQAFARVKLRRLL---TGRVEKTFKSTDSAGADVVDML 72
Db	: : : : : : : : : : : : : : : : :
351 dprpfyrlsdpvfvrvggfnphvlgltgvafqtcgtntfrnsgtidsleippg- 409	: : : : : : : : : : : : : : : : :
QY	73 TYLYNDGEFW----HPMNNTEPQLSADAKAIGDNNAKWLLDQAECIVTLWNGQPISVTTP 128
Db	: : : : : : : : : : : : : : : : : :
410 ---dngagapwndyshlvnhvtfvrrge---lsgsds-----rapmfswhr--satpt 456	: : : : : : : : : : : : : : : : : :
QY	129 NFVELELVDPDLCKGTACTGCKPATLSTGA VVKVPLFPVQIGEVTK 175
Db	: : : : : : : : : : : : : : : : : :
457 ntldperitqlpvkhtlqsg-----ttvrrgpgftg-gdlir 494	: : : : : : : : : : : : : : : : : :
RESULT	9
R84733	
ID	R84733 standard; Protein: 1148 AA.
XX	
AC	R84733;
XX	
DT	27-MAY-1996 (first entry)
XX	
XX	CryIF/cryIA(b) chimeric toxin.
DE	
KW	Delta endotoxin; crystal protein; chimeric toxin; insecticide;
KW	pesticide; Pseudomonas fluorescens; biological control agent;
KW	transgenic plant; insect resistance; disease resistance;
KW	crop improvement; protein engineering.
XX	
OS	Bacillus thuringiensis.
XX	
FH	Key Location/Qualifiers
FT	Peptide 590..601
FT	/note= "characteristic sequence of cryI toxins"
FT	Misc-difference 619
FT	/note= "E to A substitution"
FT	Misc-difference 624
FT	/note= "S to I substitution"
FT	Misc-difference 629
FT	/note= "L to I substitution"
FT	Peptide 1035..1042
FT	/note= "transition sequence of chimeric toxin"
XX	
PN	W09530753-A1.
XX	
XX	16-NOV-1995.
XX	
PF	05-MAY-1995: 95WO-US05431

XX	Synthetic.
OS	US5827514-A.
XX	27-OCT-1998.
PN	08-FEB-1996; 96US-0598305.
XX	06-DEC-1994; 94US-0349867.
XX	08-FEB-1996; 96US-0598305.
PR	(MYCO) MYCOGEN CORP.
XX	Bradfish GA, Schwab GE, Thompson M;
PI	WPI; 1998-593944/50.
DR	Composition for biological control of lepidopteran pests -
XX	comprising cells expressing two chimeric bacillus thuringiensis
PT	crystal proteins
PT	Claim 10; Column 117-124; 75pp; English.
XX	This sequence represents a Bt endotoxin which is used in a method for
PS	controlling lepidopteran pests. The method involves the use of cells
XX	that express the Bacillus thuringiensis delta-endotoxins cryIF chimeric
CC	core toxin-containing protein and CryIA(c) chimeric core toxin-containing
CC	protein in a combination that have synergistically enhanced activity,
CC	against e.g. corn earworm (Heliothis zea).
XX	Sequence 1148 AA;
SQ	

Query Match	8.0%; Score 78; DB 19; Length 1148;
Best Local Similarity	23.4%; Pred. No. 9.1;
Matches	39; Conservative 29; Mismatches 67; Indels 32; Gaps 9

QY	18 DGPVAYAEASE--FVKPGKGQAFARVKLRLL--TGTREKTEKSTDSAGADVDMNL 72
Db	: : : : : : : : : : : : : :
DB	351 dprpfvrtlsdpfvrgvgfngphylvlgirgafqgtgnhttrtfnsgrtidsideipq- 409
QY	73 TYLYNDEEPW---HPMNNETEQLSADAKAIGDNNAKWLDDQAECIVTLWNQGPISVTPP 128
Db	: : : : : : : : : : : : : :
DB	410 ---dnsgapwndyshlvnhvtivrwpge---isgsdsw----rapmfswthr--satpt 456
QY	129 NFVELEIVDTDPCLKCDTAGTGCKPATLSTGNVKKVPFLFVOIGVIK 175
Db	: : : : : : : : : : : : : : :
DB	457 ntldperitqtqipvkahtiqlsq-----ttvvrgpgftg-gdlr 494

RESULT	12
W76717	
ID	W76717 standard; Protein; 1148 AA.
XX	AC W76717;
XX	13-JAN-1999 (first entry)
DT	
XX	Alternative cryIF/cryIA(b) chimeric endotoxin protein #2.
DE	Bt toxin; lepidoptera; pest; cryIF; chimeric; core toxin; cryIA(c);
XX	synergism; plant; primer; endotoxin.
KW	Synthetic.
OS	US5827514-A.
PN	27-OCT-1998.
PD	08-FEB-1996; 96US-0598305.
PF	06-DEC-1994; 94US-0349867.
XX	

```

PR 08-FEB-1996; 96US-0598305.
XX (MYCO ) MYCOGEN CORP.
XX PA
XX XX
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX XX
XX XX WPI; 1998-593944/50.
XX DR
XX XX Composition for biological control of lepidopteran pests -
XX PT comprising cells expressing two chimeric Bacillus thuringiensis
XX PT crystal proteins
XX XX
XX PS Claim 10; Column 111-118; 75pp; English.
XX XX
XX CC This sequence represents a Bt endotoxin which is used in a method for
XX CC controlling lepidopteran pests. The method involves the use of cells
XX CC that express the Bacillus thuringiensis delta-endotoxins cryIF chimeric
XX CC core toxin-containing protein and CryIA(c) chimeric core toxin-containing
XX CC protein in a combination that have synergistically enhanced activity,
XX CC against e.g. corn earworm (Heliothis zea).
XX XX
XX SQ Sequence 1148 AA:

Query Match 8.0%; Score 78; DB 19; Length 1148;
Best Local Similarity 23.4%; Pred. No. 9.1;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 39

QY 18 DGEPIAYEASE--FVKPGKQAFARVKLRLL---TGTRVEKTFKSTDSACGADVVDNLL 72
Db 351 dprpfyrtlsdpvfrvggfnphvlgirgafqgtgnhtrtfrnsgtdsldeipqp- 409
QY 73 TYLYNDGEFW----HFMMNETPEOLSDAKAKTGDNAKLLDQAEICIVTLWNGQPSIVTTP 128
Db 410 ---dngapndyshvlnhvtfrwpg---lsqsdsw-----rapmfswthr--sacpt 456

QY 129 NFVELEIVTDPLGKGDGTAGTGKPKATLSTGAVKVPFLVQIGEVK 175
Db 457 ntldperitqlpvkahtlqsg-----ttvrgpgftg-gdlir 494

RESULT 13
ID W76707
XX W76707 standard; Protein; 1148 AA.
XX AC W76707;
XX XX
XX DT 13-JAN-1999 (first entry)
XX XX
XX DE Plasmid pMYC2244 cryIF/cryIA(b) chimeric toxin protein fragment.
XX XX
XX KW Bt toxin; lepidoptera; pest; cryIF; chimeric; core toxin; cryIA(c);
XX KW synergism; plant; primer; endotoxin.
XX OS Synthetic.
XX OS
XX PN US5827514-A.
XX XX
XX PD 27-OCT-1998.
XX XX
XX PF 08-FEB-1996; 96US-0598305.
XX XX
XX PR 06-DEC-1994; 94US-0349867.
XX PR 08-FEB-1996; 96US-0598305.
XX XX
XX PA (MYCO ) MYCOGEN CORP.
XX XX
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX XX
XX XX WPI; 1998-593944/50.
XX DR N-PSDB; V62080.
XX XX
XX PT Composition for biological control of lepidopteran pests -

```

```

PR 08-FEB-1996; 96US-0598305.
XX (MYCO ) MYCOGEN CORP.
XX PA
XX XX
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX XX
XX XX WPI; 1998-593944/50.
XX DR
XX XX Composition for biological control of lepidopteran pests -
XX PT comprising cells expressing two chimeric Bacillus thuringiensis
XX PT crystal proteins
XX XX
XX PS Claim 10; Column 111-118; 75pp; English.
XX XX
XX CC This sequence represents a Bt endotoxin which is used in a method for
XX CC controlling lepidopteran pests. The method involves the use of cells
XX CC that express the Bacillus thuringiensis delta-endotoxins cryIF chimeric
XX CC core toxin-containing protein and CryIA(c) chimeric core toxin-containing
XX CC protein in a combination that have synergistically enhanced activity,
XX CC against e.g. corn earworm (Heliothis zea).
XX XX
XX SQ Sequence 1148 AA:

Query Match 8.0%; Score 78; DB 19; Length 1148;
Best Local Similarity 23.4%; Pred. No. 9.1;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 39

QY 18 DGEPIAYEASE--FVKPGKQAFARVKLRLL---TGTRVEKTFKSTDSACGADVVDNLL 72
Db 351 dprpfyrtlsdpvfrvggfnphvlgirgvaqgtgnhtrtfrnsgtdsldeipqp- 409
QY 73 TYLYNDGEFW----HFMMNETPEOLSDAKAKTGDNAKLLDQAEICIVTLWNGQPSIVTTP 128
Db 410 ---dngapndyshvlnhvtfrwpg---lsqsdsw-----rapmfswthr--sacpt 456

QY 129 NFVELEIVTDPLGKGDGTAGTGKPKATLSTGAVKVPFLVQIGEVK 175
Db 457 ntldperitqlpvkahtlqsg-----ttvrgpgpftg-gdilr 494

RESULT 13
ID W76707
XX W76707 standard; Protein; 1148 AA.
XX AC W76707;
XX XX
XX DT 13-JAN-1999 (first entry)
XX XX
XX DE Plasmid pMYC2244 cryIF/cryIA(b) chimeric toxin protein fragment.
XX XX
XX KW Bt toxin; lepidoptera; pest; cryIF; chimeric; core toxin; cryIA(c);
XX KW synergism; plant; primer; endotoxin.
XX XX
XX OS Synthetic.
XX OS
XX PN US5827514-A.
XX XX
XX PD 27-OCT-1998.
XX XX
XX PF 08-FEB-1996; 96US-0598305.
XX XX
XX PR 06-DEC-1994; 94US-0349867.
XX PR 08-FEB-1996; 96US-0598305.
XX XX
XX PA (MYCO ) MYCOGEN CORP.
XX XX
XX PI Bradfisch GA, Schwab GE, Thompson M;
XX XX
XX XX WPI; 1998-593944/50.
XX DR N-PSDB; V62080.
XX XX
XX PT Composition for biological control of lepidopteran pests -

```



```
Qy 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL---TGTRVEKTEFKSTDSAEADVVDMNL 72
| | : | : | | : | | : | | : | | : | | : | | : | | :
Db 351 dprpfyrtlsqdvfvrgfgnphvvlgrgvafqgtgnhtrtfrnsgtidslgeippq- 409
Qy 73 TYLYNDGEFW---HFMMNETFEQLSRADAKAIGDQAKWLLDQABCIVTLWNGQPISVTPP 128
| | : | | : | | : | | : | | : | | : | | : | | :
Db 410 ---dmsgapwndyshvlnhvtfvwpge---lsgsdsw-----rapmfswthr--satpt 456
Qy 129 NFVELEIVDTPDGLKGDFTAGTGGKPKATLSTGAVVKVPLFVQIGEVIK 175
| | : | : | | : | | : | | : | | : | | : | | :
Db 457 ntldperitqiplvkahtlqsg-----ttvvrgpgftg-gdilir 494
```

Search completed: January 12, 2001, 14:58:35
Job time: 39 sec

STREET: 2421 N.W. 41st Street, Suite 200
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentLin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/598,305A
FILING DATE: 08-FEB-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,867
FILING DATE: 06-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA86.1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1148 amino acids
TYPE: amino acid

QY 129 NEVELEIVDTBGLKGDGTAGTGKGPATLSTGAVVAVPLFVQIGEVK 175
DB 457 NTIDPERITQIPLVKAHTLQSG-----TTVVRGPGFTG-GDILR 494

RESULT 9

US-08-598-305A-37
; Sequence 37, Application US/08598305A
; Patent No. 5827514
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Mark
; APPLICANT: SCHWAB, George E.
; TITLE OF INVENTION: No. 5827514el Pesticidal Compositions
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/598,305A
; FILING DATE: 08-FEB-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/349,867
; FILING DATE: 06-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA86.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-598-305A-37

Query Match 8.0%; Score 78; DB 2; Length 1148;
Best Local Similarity 23.4%; Pred. No. 3.2;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;
QY 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL---TGTTRVEKTFKSTDSAGADVVDML 72
DB 351 DPRPFYRTLSDPFVVRGGFGNPHYVLGRGVAFOQTGTHNTRFRNSGTIDSLDEIPQ- 409
QY 73 TVLYNDGEFW---HFMNNETFEQLSADAKAIGDNKWLDDQAEICVTLWNGQPISVTPP 128
DB 410 ---DMSGAPWNDYSHVLNHVTFVRWPG---ISGSDSW-----RAPMFSWTHR--SATPT 456
QY 129 NEVELEIVDTBGLKGDGTAGTGKGPATLSTGAVVAVPLFVQIGEVK 175
DB 457 NTIDPERITQIPLVKAHTLQSG-----TTVVRGPGFTG-GDILR 494

RESULT 10

US-08-598-305A-38
; Sequence 38, Application US/08598305A
; Patent No. 5827514

; GENERAL INFORMATION:
; APPLICANT: BRADFISCH, Gregory A.
; APPLICANT: THOMPSON, Mark
; APPLICANT: SCHWAB, George E.
; TITLE OF INVENTION: No. 5827514el Pesticidal Compositions
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/598,305A
; FILING DATE: 08-FEB-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/349,867
; FILING DATE: 06-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA86.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-598-305A-38

Query Match 8.0%; Score 78; DB 2; Length 1148;
Best Local Similarity 23.4%; Pred. No. 3.2;
Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;
QY 18 DGEYAVEASE--FVKPGKGQAFARVKLRLL---TGTTRVEKTFKSTDSAGADVVDML 72
DB 351 DPRPFYRTLSDPFVVRGGFGNPHYVLGRGVAFOQTGTHNTRFRNSGTIDSLDEIPQ- 409
QY 73 TVLYNDGEFW---HFMNNETFEQLSADAKAIGDNKWLDDQAEICVTLWNGQPISVTPP 128
DB 410 ---DMSGAPWNDYSHVLNHVTFVRWPG---ISGSDSW-----RAPMFSWTHR--SATPT 456
QY 129 NEVELEIVDTBGLKGDGTAGTGKGPATLSTGAVVAVPLFVQIGEVK 175
DB 457 NTIDPERITQIPLVKAHTLQSG-----TTVVRGPGFTG-GDILR 494
RESULT 11
US-08-639-923A-23
; Sequence 23, Application US/08639923A
; Patent No. 5840554
; GENERAL INFORMATION:
; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
; TITLE OF INVENTION: Pseudomonas fluorescens
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1


```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1148 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-08-639-923A-35

```

QY 18 DGEPYAVEASE--FVKPGKGQAFARVKLRRLL---TGTRVEKTFKSTDSAEGADVDMNL 72
| | :
DQ 351 DPRPFVETLSDPPEVRGGFGFNPHYVLGRGVAFOOTGTNHTRTFRNSCTIDSLDRIPPO- 409

Qy	18	DGEYAVEAGE--FVKPGKGOAFARVKLRLL---	-TCTRVKTEFKSTDSAGGVDVDMNL	72
Db	351	DPRFFYTLSDPFPVRGGFNGPHVVLGRVAGTQQTG	: : : : : : : : : : : :	409
Qy	73	TYLYNCGEFV----HFNNETFEQLSADAKAIGD	NAKLLDQAEIVTLNGQSPISVTPP	128
Db	410	---DNSGAPNDYSHVLDHVTFFVRKPE---	ISGDSW-----RAPMFSWTHR--SATPT	456
Qy	129	NFVELEIVTDTPGLKGTACTGKPKATLSTGAVV	KVLPFLQIGEVK	175
Db	457	NTIDPERITQIPLVKANTLQSG-----	TTVVRGGPFTG--GDILR	494

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match 8.0%; Score 78; DB 2; Length 1148;
Best Local Similarity 23.4%; Pred. NO. 3.2;

```

Matches 39; Conservative 29; Mismatches 67; Indels 32; Gaps 9;
Qy 18 DGEYAVEASE--FVKPGKQAFARVKLRLL---TGTRVEKTEKSTDSAEAGADVDMNL 72
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 351 DRRPFYRTLSDPVFVRGGFGNPHYVLGLRGVAFQGTGTNHTTFRNSGTIDSLDEIPQ- 409
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 73 TYLYNDGEFW-----HFMMNETFEQLSADAKAIGDNAKWLDDQAEICVTLNMGQPISVTPP 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 410 ---DNSGAPWMDYSHVLNHTFVRWPG- --ISGSDSW-----RAPMFSWTHR--SATPT 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 129 NFVELEIVDTDPGLKGTAGTGGRPATLSTGAVVKVPLFVQIGEVK 175
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 457 NTIDPERITQIPLVKAHTLQSG-----TTVVVRGPGFTG-GDILR 494
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: January 12, 2001, 14:59:00
Job time: 64 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 15:01:44 ; Search time 16.19 Seconds
(without alignments)
375.002 Million cell updates/sec

Title: US-09-322-732-2

Perfect score: 975
Sequence: 1 MATYNSNDFRAGLKIMLDGE.....QIGEVIKVDRSGEYVSRVK 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	970	99.5	187	1	EFP_ECOLI
2	744	76.3	187	1	EFP_HAEIN
3	638.5	65.5	187	1	EFP_BUCMP
4	455.5	46.7	190	1	EFP2_CHLTPR
5	449.5	46.1	192	1	EFP_AQUAE
6	445.5	45.7	190	1	EFP2_CHLPN
7	426	43.7	185	1	EFP_SNP7
8	424	43.5	187	1	EFP_SNP3
9	420	43.1	187	1	EFP_HELPY
10	417	42.8	187	1	EFP_HELPY
11	392	40.2	187	1	EFP_MYCTU
12	380	39.0	185	1	EFP_BACSU
13	377	38.7	185	1	EFP_ANASP
14	372	38.2	187	1	EFP_BRELA
15	350	35.9	185	1	EFP_TREMA
16	337	34.6	187	1	EFP_TREMA
17	300	30.8	185	1	EFP_BACFR
18	280.5	28.8	192	1	EFP_BORBU
19	252.5	25.9	188	1	EFP_RICPR
20	231	23.7	185	1	EFPI_CHLTPR
21	226	23.2	185	1	EFPI_CHLTPR
22	224	23.0	275	1	YEIP_ECOLI
23	172.5	17.7	190	1	EFP_MYCGE
24	170.5	17.5	190	1	EFP_MYCPN
25	85.5	8.8	700	1	CAN2_MOUSE
26	83	8.5	683	1	AMOH_ARTGO
27	81	8.3	1197	1	NIFJ_ANASP
28	80.5	8.3	700	1	CAN2_RAT
29	80	8.2	257	1	DNAB_MYCLE
30	79	8.1	500	1	ZIPP_DROME
31	78	8.0	1174	1	CLFA_BACTA
32	77.5	7.9	376	1	AROC_YEAST
33	76.5	7.8	513	1	YA55_SCHPO

34	76	7.8	279	1	RL2_LEPIN
35	76	7.8	440	1	OMGP_MOUSE
36	76	7.8	531	1	TRPC_PHYPR
37	76	7.8	608	1	GLMS_ECOLI
38	75	7.7	642	1	FIMB_YEAST
39	74.5	7.6	454	1	THDF_ECOLI
40	74	7.6	363	1	RF2_HELPJ
41	74	7.6	874	1	DNAB_MYCTU
42	73.5	7.5	609	1	GLMS_HAEIN
43	73	7.5	382	1	DHGY_CUCSA
44	73	7.5	638	1	SYC_HUMAN
45	72.5	7.4	299	1	Y175_HELPJ

ALIGNMENTS

RESULT 1
ID EFP_ECOLI STANDARD; PRT; 187 AA.
AC P33398;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
GN EFP.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-32; 70-79 AND 85-116.
RX MEDLINE=92066471; PubMed=1956781;
RA Aoki H., Adams S.-L., Chung D.-G., Yaguchi M., Chuang S.-E.,
RA Ganoza M.C.;
RT "Cloning, sequencing and overexpression of the gene for prokaryotic
RT factor EF-P involved in peptide bond synthesis.";
RL Nucleic Acids Res. 19:6215-6220(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Hurland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [3]
RP SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [4]
RP CHARACTERIZATION.
RC MEDLINE=97338480; PubMed=9195040;
RX Aoki H., Adams S.-L., Turner M.A., Ganoza M.C.;
RT "Molecular characterization of the prokaryotic efp gene product
RT involved in a peptidyltransferase reaction.";
RL Biochimie 79:7-11(1997).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC 70S RIBOSOMES AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC TRANSLATION IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE.
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch). CC
CC
CC EMBL; AF003957; AAC04239.1; -. DR
CC PFAM; PF01132; EEP; 1. DR
CC PROSITE; PS01275; EFP; 1. DR
CC Protein biosynthesis; Elongation factor. KW
CC NON_TER 187 187 FT
CC SEQUENCE 187 AA; 211535 MW; 9C803DCA4EEEA025 CRC64; SQ

Query Match	65.5%	Score 638.5	DB 1	Length 187
Best Local Similarity	64.7%	Pred. No. 2.6e-50		
Matches 121	Conservative	27	Mismatches 38	Indels 1
Gaps				
QY	1	MATVYSNDRFAGLKIMLDGEPYAVEASEFEVKPGKGAFARVKLRLLLTGTVEKTKSTD	60	
		: : : : : : : : : : : : : :		
Db	1	MRVHSHNFSGCKKIIFENEPYLIESSEFEVKPGKGAFVRVKRKLRLTQLRIETKFKSTD	60	
QY	61	SABGADVVMNLIVLYNDGFHFHMNNTEPEQLSADAKAIGDNAKWLLDAQECITVTLWNG	120	
		: : : : : : : : : :		
Db	61	YLEIADVIEHTCLYLYNDGFHWFINNTFEELSVEKKIKGIQKWLLEQDTCTIVTLWN	120	
QY	121	OPTSVTPPNFVELIEVDTPDLGLGDTAGTG-GKPATILSTGAUVKVPFLVQIGEVIKVDTTR	179	
		: : : : : : : :		
Db	121	KPLIVTPNTPEIKVIEDATLKGDTINTGVTKLAILSTGAIVRPVLFIQVSLIKVDTTR	180	
QY	180	SGEVSR	186	
Db	181	SGEVSR	187	
RESULT	4			
EFP2_CHLTR		STANDARD;	PRT;	190 AA.
ID	EFP2_CHLTR			
AC	O84757			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	ELONGATION FACTOR P 2 (EF-P 2).			
GN	EFP2 OR CT752.			
OS	Chlamydia trachomatis.			
Bacteria:	Chlamydiales; Chlamydiaceae; Chlamydia.			
[]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=D/UW-3/CX;			
RC	MEDLINE=99000809; PubMed=9784136;			
RX	Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,			
RA	McIntire W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,			
RA	Davis R.W.;			
RA	"Genome sequence of an obligate intracellular pathogen of humans:			
RT	Chlamydia trachomatis."			
RT	Science 282:754-759(1998).			
RL	-1- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT			
CC	TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED			
CC	70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING			
CC	THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING			
CC	THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE			
CC	(BY SIMILARITY).			
CC	-1- PATHWAY: PROTEIN BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.			
CC	-1- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.			

```
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE001347; AAC68347.1; -.
DR PRAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 190 AA; 21520 MW; 75F0D7EBEEFC20D7 CRC64;

Query Match          46.7%; Score 455.5; DB 1; Length 190;
Best Local Similarity 48.9%; Pred. No. 6.1e-34;
Matches 92; Conservative 36; Mismatches 59; Indels 1; Gaps 1;

QY 1 MATVYSNDKIMLDGEPYAYEASEFEVKPGKGQAFARVKLRRLLTGTVEKTFKSTD 60
DB 1 MVRVSTSEFRVGLRVKIDGQPVIQLQDFVKPGKGQAFNRKVNFILTGRVIEKTFKSGE 60
      : ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 SAEGADVVDMLTYLYNDGEFWHFMMNETPQLSADAKAIGDNKWLDDQAECITVLWNG 120
DB 1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 QPISVTTPNPEVEIIVDPDLGKDTA-GTCCKKATLSTGTGAWKVPFLPVQIGEVIKKVDTR 179
DB 1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 DVISVEPIPFMELTAETAPGVRGDTASGRVLKPATTNTGAKINMPVFIEEGEVKKVDTR 180
DB 1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 180 SGEVSVSR 187
DB 1 TGSYESRV 188

RESULT 5
EFP_AQUAE STANDARD; PRT; 192 AA.
AC O67376;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EF-P).
EN EFP OR AQ_1364.
OS Aquifex aeolicus.
SC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RC STRAIN-VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
   aeolicus";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000736; AAC07331.1; -.
DR INTERPRO; IPR001059;-.
```

```
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor.
SQ SEQUENCE 192 AA; 21807 MW; 9B651E676507CAA0 CRC64;

Query Match 46.1%; Score 449.5; DB 1; Length 192;
Best Local Similarity 48.7%; Pred. No. 2.1e-33;
Matches 92; Conservative 31; Mismatches 65; Indels 1; Gaps 1;

QY 1 MATYYS-NDPRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTGRVEKTEKST 59
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1 MATEIDINRQKDFIEHKGEPRVLDYEHVKPGKGQAFARVKLRLLTGTGRVEKTEKST 60
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

QY 60 DSABGADVDMNLTYLYNDGEFHFMMNNEFTEQLSADAKAIGDNNAKLLDQAEICIVTLWN 119
   || : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 61 DRIPLADEQVYATYSYNDGENYFMNTQYDMIAVPKEIEEAKFLKEGMEVIVFLYK 120

QY 120 QPISVTPPNFVELEIVDTPDGLKGDGTAGTGKGPATLSTGAVVKVPLFVQIGEVIKVDTR 179
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 121 GQPTIGIELPKHVELQVVEFTEPAFGKDTQAGGTPAKLETGAVIQVPPFVKEGDVIKVDTR 180

QY 180 SGEYVSRVK 188
   : || || |||
Db 181 TGSYVERVK 189

RESULT 6
EFP2_CHLPN STANDARD; PRT; 190 AA.
AC Q9Z711;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE ELONGATION FACTOR P 2 (EFP 2).
GN EFP2 OR CPN0895.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: A501670; AAD19033.1; -.
CC DR A501670; EFP; 1.
CC DR PROSITE; PS01275; EFP; 1.
CC KW Protein biosynthesis; Elongation factor.
CC SQ SEQUENCE 190 AA; 21477 MW; 1A245E50BDAA340C CRC64;

Query Match 45.7%; Score 445.5; DB 1; Length 190;
Best Local Similarity 46.3%; Pred. No. 4.8e-33;
Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;
```

```
QY 1 MATYYSNDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTGRVEKTEKST 60
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| :
Db 1 MVRVSTSEPRVGLRIEIDGQPYLLQNDQFVKPGKGQAFNRKVKNFELTGRVIERVTKSGE 60

QY 61 SAGGADVDMNLTYLYNDGEFHFMMNNEFTEQLSADAKAIGDNNAKLLDQAEICIVTLWN 120
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 61 SVETADIVERSMRLLYTDQEGATFMDDETPEQEVFWEKLENIROWLLEDITVTLVLYNG 120

QY 121 QPISVTPPNFVELEIVDTPDGLKGDGTAGTGKGPATLSTGAVVKVPLFVQIGEVIKVDTR 179
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| :
Db 121 DVVAEPPFPMELISIAETAPGVRGDTASGRVLPKPAVTNTGAKIMVPIFDIDEGELVKVDTR 180

QY 180 SGEYVSRV 187
   : || |||
Db 181 TGSYESRV 188

RESULT 7
EFP_SYNPF7 STANDARD; PRT; 185 AA.
AC Q54760;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ELONGATION FACTOR P (EFP-P).
GN EFP.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
RN [1]
RP SEQUENCE FROM N.A.
RA Phung L.T., Haselkorn R.;
RT "Genes encoding biotin carboxyl carrier protein and elongation
RT factor P from cyanobacterium Synecococcus sp. PCC 7942.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U59235; AAB82025.1; -.
CC DR U59235; AAB82025.1; -.
CC DR INTERPRO; IPR001059; -.
CC DR PFAM; PF01132; EFP; 1.
CC DR PROSITE; PS01275; EFP; 1.
CC KW Protein biosynthesis; Elongation factor.
CC SQ SEQUENCE 185 AA; 20368 MW; BB6880A3D96CF06C CRC64;

Query Match 43.7%; Score 426; DB 1; Length 185;
Best Local Similarity 48.6%; Pred. No. 2.5e-31;
Matches 88; Conservative 25; Mismatches 68; Indels 0; Gaps 0;

QY 6 SDFRAGLKIMLDGEPYAVEASEFVKPGKGQAFARVKLRLLTGTGRVEKTEKSTDSAGA 65
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 4 SDFRTGTTIEIDGAVRVVVEFLHVKPGKGSFVTRTKLNAKTGNVVEKTRFAGTVPQA 63

QY 66 DVVDMNLTYLYNDGEFHFMMNNEFTEQLSADAKAIGDNNAKLLDQAEICIVTLWNQPIV 125
   : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| :
Db 64 VLEKSTLQTYKDGDDDFVFMDETYEEGRLTATIGDRVYKLEKMEANVTWNGQVIEV 123
```


[illegible]

Db 179 GSYLGRV 185

RESULT 12

EFP_BACSU

ID EFP_BACSU STANDARD; PRT; 185 AA.

AC P49778;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE ELONGATION FACTOR P (EF-P).

GN EFP.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-168 / JH642;

RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,

RA Sato T., Takeuchi M.;

RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 139-185 FROM N.A.

RC STRAIN-168 / JH642;

RA Guerout-Fleury A.M., Gonzy-Treboul G., Stragler P.;

RA Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT

CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED

CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING

CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING

CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE

CC (BY SIMILARITY).

CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D84432; BAA12558.1; -.

DR EMBL; U35252; AAA76718.1; -.

DR EMBL; Z99116; CAB14376.1; -.

DR SUBTILIST; BG11460; EFP.

DR INTERPRO; IPR001059; -.

DR PFAM; PF01132; EFP; 1.

DR PROSITE; PS01275; EFP; 1.

KW Protein biosynthesis; Elongation factor.

SQ SEQUENCE 185 AA; 20454 MW; 60C3B9BF90D2DAB8 CRC64;

Query Match

Best Local Similarity 39.0%; Score 380; DB 1; Length 185;

Matches 78; Conservative 30; Mismatches 72; Indels 0; Gaps 0;

Qy 7 NDFRAGLKIMLDGPEYAVESEFVKPGKQAFARVKLRLLTGTTRVEKTFKSTDSAGAD 66

Db 5 NDFRTGLTIDVDGGIWRVDFQHKVPGKGAFAVRKSLRLTGTATQETFRAGEKVAQAQ 64

Qy 67 VVDNMNLTLYNDGEFHFHNNNETFEQLSADAKAIGDGNKWLDDQACIVTLWNGQPSIVT 126

Db 65 IETKTMQYLYANGDQHFMDTSSEYQELSATQIEELKYLLENKSVHIMMYQDETLLGIE 124

Qy 127 PPNFVELEIVDTDPGLKGDGTAGTGKGPATLSTGAVKVPLEFVQIGEVKIVDTRSGEYVSR 186

Db 125 LPNIVELKVVETEPGKIGKDGTSAGGKTPAKTETGLVNVNPFVFNEDGDTLLVWNTSDGSYVSR 184

RESULT 13

EFP_ANASP

ID EFP_ANASP STANDARD; PRT; 185 AA.

AC O44247;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE ELONGATION FACTOR P (EF-P).

GN EFP.

OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-93352435; PubMed-8102363;

RA Gornicki P., Scappino L.A., Haselkorn R.;

RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena

RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier

RT protein.";

RL J. Bacteriol. 175:5268-5272(1993).

CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT

CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED

CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING

CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING

CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE

CC (BY SIMILARITY).

CC -!- PATHWAY: PROTEIN BIOSYNTHESIS.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L14863; AAA74627.1; -.

DR INTERPRO; IPR001059; -.

DR PFAM; PF01132; EFP; 1.

DR PROSITE; PS01275; EFP; 1.

KW Protein biosynthesis; Elongation factor.

SQ SEQUENCE 185 AA; 20496 MW; F5F32A12CF106182 CRC64;

Query Match

Best Local Similarity 38.7%; Score 377; DB 1; Length 185;

Matches 78; Conservative 33; Mismatches 70; Indels 0; Gaps 0;

Qy 6 SNDFRAGLKIMLDGPEYAVESEFVKPGKQAFARVKLRLLTGTTRVEKTFKSTDSAGA 65

Db 4 SNDFRPGSVIVLDGSMVRVIDFLHVKPGKGSFAFVTTTLKNVQSKVLEKTFRAGETVPOA 63

Qy 66 DVVDMNLTLYNDGEFHFHNNNETFEQLSADAKAIGDGNKWLDDQACIVTLWNGQPSIV 125

Db 64 TLEKITMQHTYKEGDEFVMDMESYEGRLSAAQIGDRVKYLKEGMEVNVIRWGEQVLEV 123

Qy 126 TPNFVELEIVDTDPGLKGDGTAGTGKGPATLSTGAVKVPLEFVQIGEVKIVDTRSGEYVS 185

Db 124 ELANSVVLEVIQTDPGVKGDGTATGTPAIVETGATVNVPLFISQGERIKIDTRDDRYLG 183

Qy 186 R 186

Db 184 R 184

RESULT 14

EFP_BRELA

ID EFP_BRELA STANDARD; PRT; 187 AA.

AC Q45288;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2001, 15:00:08 ; Search time 39.67 Seconds
(without alignments)
555.460 Million cell updates/sec

Title: US-09-322-732-2
Perfect score: 975
Sequence: 1 MATYNSNDRFRAGLKIIMLDGE.....QIGEVIKVTRSGEYVSRVK 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	803	82.4	161	2	O69281 citrobacter
2	750	76.9	188	2	O9KNS1 vibrio chol
3	638	65.4	193	2	O9PAM3 O9pam3 xylella fas
4	450.5	46.2	190	2	O9PLH1 O9plh1 chlamydia m
5	445.5	45.7	190	2	O9TOB4 O9tob4 chlamydia p
6	438	44.9	189	2	O9PHW3 O9phw3 campylobact
7	414	42.5	185	2	O9K951 O9k951 bacillus ha
8	405	41.5	184	10	O9M9I7 O9m9i7 arabisdopsis
9	378.5	38.8	188	2	O9KXQ9 O9kxq9 streptomyc
10	358	36.7	185	2	O9RY32 O9ry32 deinococcus
11	317	32.5	136	2	O9S3F5 O9s3f5 helicobacte
12	315	32.3	136	2	O9S3F9 O9s3f9 helicobacte
13	315	32.3	136	2	O9S3F6 O9s3f6 helicobacte
14	315	32.3	136	2	O9R2M0 O9r2m0 helicobacte
15	312	32.0	136	2	O9S3F7 O9s3f7 helicobacte
16	312	32.0	136	2	O9R3F4 O9r3f4 helicobacte
17	311	31.9	136	2	O9S3F8 O9s3f8 helicobacte
18	310	31.8	136	2	O9S3G0 O9s3g0 helicobacte
19	291.5	29.9	186	2	O9JU02 O9juu2 neisseria m

20	290.5	29.8	186	2	O9JZ08 O9jz08 neisseria m
21	283	29.0	187	2	O9RNJ9 O9rnj9 zymomonas m
22	235.5	24.2	199	2	O9KSP7 O9ksp7 vibrio chol
23	226.5	23.2	189	2	O9PBE1 O9pbe1 xylella fas
24	226	23.2	185	2	O9JQ76 O9jq76 chlamydia p
25	221	22.7	185	2	O9PKR6 O9pkr6 chlamydia m
26	200	20.5	188	2	O9PQJ3 O9pqj3 ureaplasma
27	157	16.1	109	2	P82459 P82459 streptococc
28	144.5	14.8	212	10	O9STQ4 O9stq4 arabisdopsis
29	87	8.9	161	2	O45051 O45051 borrelia bu
30	86.5	8.9	700	11	O54843 O54843 mus musculu
31	86	8.8	289	2	O46146 O46146 chromohalob
32	84.5	8.7	408	2	O9XIR3 O9xir3 thermotoga
33	84	8.6	307	1	O9UXK6 O9uxk6 sulfobus
34	83	8.5	161	2	O51614 O51614 borrelia bu
35	83	8.5	383	2	O9RV55 O9rv55 deinococcus
36	83	8.5	1933	1	O9YEF6 O9yef6 aeropyrum p
37	80.5	8.3	997	1	O9UY74 O9uy74 pyrococcus
38	80	8.2	604	2	O53124 O53124 mycobacteri
39	79	8.1	245	5	O17439 O17439 boltonia vi
40	78.5	8.1	476	5	O16663 O16663 caenorhabdi
41	78.5	8.1	549	2	O82975 O82975 bacillus sp
42	78	8.0	210	2	O9KB07 O9kb07 bacillus ha
43	77.5	7.9	1430	5	O97381 O97381 artemia sal
44	77	7.9	303	2	O9Z9P7 O9z9p7 bacillus ha
45	77	7.9	372	2	O9KFG6 O9kfg6 bacillus ha

ALIGNMENTS

RESULT 1

O69281 PRELIMINARY; PRT; 161 AA.

AC O69281;

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE ELONGATION FACTOR-P HOMOLOG (FRAGMENT).

OS Citrobacter freundii.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Citrobacter.

OX NCBI_TaxID=546;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OS60;

EX MEDLINE=96032747; PubMed=7559452;

RA Bishop R.E., Penfold S.S., Frost L.S., Holtje J.V., Weiner J.H.;

RT "Stationary phase expression of a novel Escherichia coli outer

RT membrane lipoprotein and its relationship with mammalian

RT apolipoprotein D. Implications for the origin of lipocalins.";

RL J. Biol. Chem. 270:23097-23103(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OS60;

RA Bishop R.E., Leski B.K., Hodges R.S., Kay C.M., Weiner J.H.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: U21727; AAC46461.1; -

DR INTERPRO: IPR001059; -

DR PFAM: PF01132; EFP; 1.

DR PROSITE: PS01275; EFP; 1.

KW Elongation factor.

FT NON_TER 1

SQ SEQUENCE 161 AA; 17525 MW; D475D5420DF0BEEC4 CRC64;

Query Match 82.4%; Score 803; DB 2; Length 161;
Best Local Similarity 94.4%; Pred. No. 1.4e-65;
Matches 152; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 28 EPVKPGKGAFAFARVKLRLLTCTRVEKFKSTDSAGADVDMNLTYLYNDGEFHFMMN 87
|||||

DB 1 EPVKPGKGAFAFARVKLRLLTCTRVEKFKSTDSAGADVDMNLTYLYSDGEFHFMMN 60
|||||


```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman W., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O.,
RA Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Bass S., Linher K., Weidman W., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE002280; AAF39011.1; -.
DR TIGR; TC0133; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
DR SEQUENCE 190 AA; 21550 MW; C5920D53944CD9BD CRC64;

Query Match 46.2%; Score 450.5; DB 2; Length 190;
Best Local Similarity 48.4%; Pred. No. 2e-33;
Matches 91; Conservative 36; Mismatches 60; Indels 1; Gaps 1;

QY 1 MATYNSDNFRAGLKIMLDGEPYAVEASEFVKPGKQAFARVKLRLLTGTRVEKTFKSTD 60
DQ 1 MVRVSTSEFRVGLRIEIDGQPVYLQNDVFVKPGKQAFNRKVKNFGLTGRVIERKSGE 60
QY 61 SAEGADVDMNLTLYNDGEFHFNNETFEOLSADAKAIGDNKWLDDQAEICVTLWNG 120
DQ 61 SIETADVREQMRLTYDOEGATFMDDETFEQELIFWFKLENIRQWLLDVTYTLVRYNG 120
QY 121 QPISVTPPNFVELEIVDFDPGLKGDTA-GTGGKPATLSTGAVVKVPLFVQIGEVKVDTR 179
DQ 121 DVISVEPPIFMELSIETAETAPGVGDGTASGRVLKPAVTNTGAKIMVPFIPIFEGELVKVDTR 180
QY 180 SGEYVS RV 187
DQ 181 TGSYESRV 188

RESULT 5
QYQB4 PRELIMINARY; PRT; 190 AA.
AC QYQB4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE ELONGATION FACTOR P (TRANSLATION ELONGATION FACTOR P).
GN EFP_2 OR CP0971.
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
```

```
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman W., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AP002548; BAA99103.1; -.
DR EMBL; AE002255; AAF38751.1; -.
DR TIGR; CP0971; -.
KW Elongation factor.
SQ SEQUENCE 190 AA; 21477 MW; 1A245E50BDAA340C CRC64;

Query Match 45.7%; Score 445.5; DB 2; Length 190;
Best Local Similarity 46.3%; Pred. No. 5.6e-33;
Matches 87; Conservative 42; Mismatches 58; Indels 1; Gaps 1;

QY 1 MATYNSDNFRAGLKIMLDGEPYAVEASEFVKPGKQAFARVKLRLLTGTRVEKTFKSTD 60
DQ 1 MVRVSTSEFRVGLRIEIDGQPVYLQNDVFVKPGKQAFNRKVKNFGLTGRVIERKSGE 60
QY 61 SAEGADVDMNLTLYNDGEFHFNNETFEOLSADAKAIGDNKWLDDQAEICVTLWNG 120
DQ 61 SVETADIVERSMRLTYDOEGATFMDDETFEQEVVFWKLENIRQWLLDVTYTLVLYNG 120
QY 121 QPISVTPPNFVELEIVDFDPGLKGDTA-GTGGKPATLSTGAVVKVPLFVQIGEVKVDTR 179
DQ 121 DVAVPEPPIFMELSIETAETAPGVGDGTASGRVLKPAVTNTGAKIMVPFIPIFEGELVKVDTR 180
QY 180 SGEYVS RV 187
DQ 181 TGSYESRV 188

RESULT 6
QYQB4 PRELIMINARY; PRT; 189 AA.
AC QYQB4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ELONGATION FACTOR P.
GN EFP.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jajels K., Karlyshev A., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL; AL139075; CAB75187.1; -.
DR INTERPRO; IPR001059; -.
DR INTERPRO; IPR001884; -.
DR PFAM; PF01132; EFP; 1.
DR PFAM; PF01287; eif-5a; 1.
SQ SEQUENCE 189 AA; 21098 MW; COBAE2CC7D220300 CRC64;

Query Match 44.9%; Score 438; DB 2; Length 189;
Best Local Similarity 45.7%; Pred. No. 2.7e-32;
Matches 85; Conservative 33; Mismatches 68; Indels 0; Gaps 0;
```



```
Db 4 TNDLRNGLVLKLEGGQLWSVEFHVKPGKGFPAFVRTKLKNVLSGKVVYDKTFNAGVKVET 63
QY 65 ADVVDMNLTLYNDGFEHFHMNNETFEQLSADAKAIGNAKWLLDQAEICVTLMNGQPIIS 124
Db 64 ATVDKRDQFSDMGDYFVFMDETYDQIMDRKVVGDAAANFLVGFEGFATVAHQHEGVLF 123
QY 125 VTPPNFVELEIVDTPGLKGDGTAGTGGKPAATLSTGAVVKVPLFVQIGEVKVDTRSGEYV 184
Db 124 VFLPAAVELTIQETEPGVQGGDRSTGCTKTPATLETGHQINVPLFITTGKIKVDTRTSYL 183
QY 185 SRV 187
Db 184 GRV 186

RESULT 10
Q9RY32 PRELIMINARY; PRT; 185 AA.
AC Q9RY32;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P.
GN DR0119.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Matharova J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makareva K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans RL."
RL Science 286:1571-1577(1999).
DR EMBL; AE001874; AAF09709.1; -.
DR TIGR; DR0119; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Elongation factor.
SQ SEQUENCE 185 AA; 20475 MW; 893FD0850AC102D3 CRC64;

Query Match 36.7%; Score 358; DB 2; Length 185;
Best Local Similarity 40.8%; Pred. No. 4,8e-25;
Matches 73; Conservative 38; Mismatches 68; Indels 0; Gaps 0;

QY 8 DFRAGLKIMLDGEPXAVEASEFVKGQAFARVLRLLLTGTRVEKFTKSDSAEGADV 67
Db 6 ELRNGTKVQMDGLWECLDYSLKMGKGGCAKVVTKFRNMGSGSIVDRITFNSTEKLQDIIV 65
QY 68 VDMNLTLYNDGFEHFHMNNETFEQLSADAKAIGNAKWLLDQAEICVTLMNGQPIISVTP 127
Db 66 EGKMKQLYPDGDYVFMDETYDQIMDRKVVGDAAANFLVGFEGFATVAHQHEGVLF 123
QY 128 PNFVELEIVDTPGLKGDGTAGTGGKPAATLSTGAVVKVPLFVQIGEVKVDTRSGEYV 186
Db 126 PNVILKITQDPGVGRGDTVSGGTRKTPATLETGAVVQVPLFVEQGGTDVKVDTRGYLSR 184

RESULT 11
Q9S3F5 PRELIMINARY; PRT; 136 AA.
AC Q9S3F5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
```

```
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPI;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239624; CAB37735.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15008 MW; B317AF0A270ADA48 CRC64;

Query Match 32.5%; Score 317; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 1,7e-21;
Matches 61; Conservative 28; Mismatches 46; Indels 0; Gaps 0;

QY 28 EFVPGKQAFARVLRLLLTGTRVEKFTKSDSAEGADVDMNLTLYNDGFEHFHMNN 87
Db 2 QHVPGKGAARVRAKISFLDKGVIEKTFHAGDKCEKSNLVEKTMQYLYHGDYTFQMDI 61
QY 88 ETEFOLSADAKAIGNAKWLLDQAEICVTLMNGQPIISVTPNFVELEIVDTPGLKGDFTA 147
Db 62 ESYEIOIALNDQSGEASKWMLDGMQVQLLNDKRAISVDVFOVWALKIVETAPNFKGDTIS 121
QY 148 GTGKGPATLSTGAVV 162
Db 122 SASKKPALETCAV 136

RESULT 12
Q9S3F9 PRELIMINARY; PRT; 136 AA.
AC Q9S3F9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#12;
RA Achtman M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
RA Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
RT different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239608; CAB37719.1; -.
DR INTERPRO; IPR001059; -.
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15049 MW; FCCA2867F2878289 CRC64;

Query Match 32.3%; Score 315; DB 2; Length 136;
```

```

Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
: ||||| || | : | : |||| | : | : | : | : | : | : | : | : | : | :
Db 2 QHVPGKGAAFVTRIKISFLDGVKIEKTFHAGDKCEEPNLVEKTMQYLYHGDGTQFMDI 61
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
QY 88 ETEFOLSADAKAIGNAKWLDDQACIVTLWNGQIPISVTPPNFVELEIVTDPLKGGDTA 147
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
Db 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
QY 148 GTGGKPATLSTGAVV 162
: ||||| |||||
Db 122 SASKKPATLETGAVV 136
: ||||| |||||

RESULT 13
Q9S3F6 PRELIMINARY; PRT; 136 AA.
AC Q9S3F6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239613; CAB37724.1; -
DR EMBL; AJ239614; CAB37725.1; -
DR EMBL; AJ239615; CAB37726.1; -
DR EMBL; AJ239616; CAB37727.1; -
DR EMBL; AJ239617; CAB37728.1; -
DR EMBL; AJ239618; CAB37729.1; -
DR EMBL; AJ239619; CAB37730.1; -
DR EMBL; AJ239620; CAB37731.1; -
DR EMBL; AJ239621; CAB37732.1; -
DR EMBL; AJ239622; CAB37733.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15018 MW; FCC35872F28FA29C CRC64;

Query Match 32.3%; Score 315; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
: ||||| || | : | : |||| | : | : | : | : | : | : | : | : | : | :
Db 2 QHVPGKGAAFVRAKISFLDGVKIEKTFHAGDKCEEPNLVEKTMQYLYHGDGTQFMDI 61
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
QY 88 ETEFOLSADAKAIGNAKWLDDQACIVTLWNGQIPISVTPPNFVELEIVTDPLKGGDTA 147
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
Db 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
QY 148 GTGGKPATLSTGAVV 162
: ||||| |||||
Db 122 SASKKPATLETGAVV 136
: ||||| |||||

RESULT 14
Q9R2M0 PRELIMINARY; PRT; 136 AA.
AC Q9R2M0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).

```

```

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239625; CAB37736.1; -
DR EMBL; AJ239613; CAB37724.1; -
DR EMBL; AJ239614; CAB37725.1; -
DR EMBL; AJ239615; CAB37726.1; -
DR EMBL; AJ239616; CAB37727.1; -
DR EMBL; AJ239617; CAB37728.1; -
DR EMBL; AJ239618; CAB37729.1; -
DR EMBL; AJ239619; CAB37730.1; -
DR EMBL; AJ239620; CAB37731.1; -
DR EMBL; AJ239621; CAB37732.1; -
DR EMBL; AJ239622; CAB37733.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15019 MW; FCC35872F287829C CRC64;

Query Match 32.3%; Score 315; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 2.6e-21;
Matches 61; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 28 EFVPGKGQAFARVKLRLLTGTTRVEKTFKSTDSAGADVDDMNLTYLYNDGEFHFMMN 87
: ||||| || | : | : |||| | : | : | : | : | : | : | : | : | : | :
Db 2 QHVPGKGAAFVRAKISFLDGVKIEKTFHAGDKCEEPNLVEKTMQYLYHGDGTQFMDI 61
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
QY 88 ETEFOLSADAKAIGNAKWLDDQACIVTLWNGQIPISVTPPNFVELEIVTDPLKGGDTA 147
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
Db 62 ESYEQIALNDSQVGEASKWLDGMQVQLLHNDKKAISVDVQVVALKIVETAPNFKGDT 121
: ||||| : | : | : |||| | : | : | : | : | : | : | : | : | : | :
QY 148 GTGGKPATLSTGAVV 162
: ||||| |||||
Db 122 SASKKPATLETGAVV 136
: ||||| |||||

RESULT 15
Q9S3F7 PRELIMINARY; PRT; 136 AA.
AC Q9S3F7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ELONGATION FACTOR P (FRAGMENT).
GN EFP.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N6;
RA Achtnan M., Azuma T., Berg D.E., Ito Y., Morelli G., Pan Z.J.,
Suerbaum S., Thompson S., van der Ende A., van Doorn L.J.;
RT "Recombination and clonal groupings within Helicobacter pylori from
different geographic regions.";
RL Mol. Microbiol. 32:459-470(1999).
DR EMBL; AJ239611; CAB37722.1; -
DR INTERPRO; IPR001059; -
DR PFAM; PF01132; EFP; 1.
KW Elongation factor.
FT NON_TER 1
FT NON_TER 1

```

FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15018 MW; FCCDB89CF287829C CRC64;

Query Match 32.0%; Score 312; DB 2; Length 136;
Best Local Similarity 45.2%; Pred. No. 4.8e-21;
Matches 61; Conservative 26; Mismatches 48; Indels 0; Gaps 0;
Qy 28 EFVKPGKQAFARVKLRRLTGTTRVEKTFKSTDSAEGADVDDMNLTYLYNDGEFWHFMMN 87
 : ||||| || | : : | : |||| : | : : | : |||| : : || :
Db 2 QHVKPGKGAAPVRAKIKSFLDGKVIKTFHAGDKCEEPNLEKTMQVLYLHGDITYQFMDI 61
Qy 88 ETFEQLSADAKAIGDNAKWLDDQAEICVTLWNGOPISVTPPNFVELEIVDTPGLKGDTA 147
 | : || : : : | : || : || : | : | : ||| | | : ||| |
Db 62 ESYEQIALNDSQVGEASKWMLDGMQVQLLHNDKKAISVDVPQVVALKIVKTAPNFKGDT 121
Qy 148 GTGGKPATLSTGAVV 162
 ||||| |||||
Db 122 SASKKPATLETGAVV 136

Search completed: January 12, 2001, 15:00:10
Job time: 130 sec

